Self-Regulating Cars: Automating Traffic Control in Free Flow Road Networks

Ankit Bhardwaj[1], **Rohail Asim**[1], **Sachin Kumar Chauhan**[2], **Yasir Zaki**[1], **Lakshmi Subramanian**[1]

[1]Department of Computer Science, New York University, [2]Department of Computer Science, IIT Delhi
bhardwaj.ankit@nyu.edu, rohail.asim@nyu.edu, csz188012@cse.iitd.ac.in, yasir.zaki@nyu.edu, lakshmi@nyu.edu

Abstract

Free-flow road networks, such as suburban highways, are increasingly experiencing traffic congestion due to growing commuter inflow and limited infrastructure. Traditional control mechanisms—traffic signals or local heuristics—are ineffective or infeasible in these high-speed, signal-free environments. We introduce self-regulating cars, a reinforcement learning-based traffic control protocol that dynamically modulates vehicle speeds to optimize throughput and prevent congestion, without requiring new physical infrastructure. Our approach integrates classical traffic flow theory, gap acceptance models, and microscopic simulation into a physics-informed RL framework. By abstracting roads into super-segments, the agent captures emergent flow dynamics and learns robust speed modulation policies from instantaneous traffic observations. Evaluated in the high-fidelity PTV Vissim simulator on a real-world highway network, our method improves total throughput by 5%, reduces average delay by 13%, and decreases total stops by 3% compared to the no-control setting. It also achieves smoother, congestion-resistant flow while generalizing across varied traffic patterns—demonstrating its potential for scalable, ML-driven traffic management.

Code and Simulation Files: https://github.com/ankitbha/Self-Regulating-Cars

Introduction

Traffic jams lead to significant economic and productivity losses globally (Badger 2013; Downie 2008), disproportionately affecting developing regions (Bhardwaj et al. 2023). The growing volume of traffic now increasingly affects previously uncongested, high-speed freeways due to suburban sprawl and commuter traffic flows (Gordon et al. 1996). The concept of managing congestion not being a primary consideration in the initial design, these freeways are largely un-signalized. The straightforward solution of infrastructure expansion (building new roads) and modification (installing signals) remains prohibitively expensive, leading to perennial jams. Leveraging the rise of AI-driven vehicles, we introduce *self-regulating cars*, an automated traffic management system where smart vehicles adjust their speeds using a centralized protocol. This system enables dynamic, infrastructure-free congestion control by coordinating individual vehicle speeds to maintain high throughput and prevent traffic jams.

We frame freeway congestion control as a centralized reinforcement learning (RL) problem, where an agent dynamically sets segment-level *desired speeds* to optimize traffic flow. Unlike traditional traffic signaling protocols, which rely on distributed, reactive heuristics at intersections, our method learns a global policy that coordinates across the network while accounting for flow-density dynamics and stochastic inflow and driving behaviors. We incorporate domain knowledge from traffic flow theory, gap acceptance models, and car-following dynamics into the RL design: key traffic metrics (e.g., density, speed, inflow) form the state space, and the reward penalizes congestion while promoting high-speed flow. The agent operates in a discretized action space to support human-in-the-loop deployment and actuation robustness. We train and evaluate our system using the PTV Vissim microscopic simulator on a real-world free-flow road network in Mainz, Germany. Compared to baseline signaling protocols, our approach consistently achieves higher throughput, smoother flow, and greater resilience under varied traffic conditions.

A central challenge in our setting is maintaining smooth, uninterrupted flow in the presence of stochastic traffic inflows and dynamic merging behavior. Unlike signalized traffic control, where stop-and-go behavior is imposed by design, free-flow networks are vulnerable to spontaneous congestion waves that emerge from local interactions, even without explicit bottlenecks. Our RL agent must therefore learn a delicate balance: modulating speeds enough to prevent congestion, but not so aggressively as to induce braking cascades or shockwaves. To enable this, we aggregate vehicle dynamics over larger *super-segments* to capture emergent flow behaviors while smoothing over short-term fluctuations. The discrete action space reflects real-world deployability constraints, enabling implementation via in-vehicle apps or interfaces, while the use of instantaneous state representations, instead of temporal features, supports generalization across varied and stochastic traffic patterns. This design ensures the agent learns policies that are physically grounded and robust to input shifts, rather than overfitting to specific temporal trajectories.

In this paper, we address these challenges through the following key contributions:

- We formulate freeway traffic flow regulation as a centralized reinforcement learning problem aimed at maximizing

network throughput while maintaining smooth, congestion-free operation in stochastic, free-flow settings.
- We design a physics-informed RL framework that incorporates traffic-theoretic insights into both state and reward structures, enabling robust and interpretable policy learning without temporal input dependencies.
- We introduce a super-segment abstraction to capture emergent flow dynamics and mitigate local noise, and leverage a discretized action space to support human-in-the-loop deployment via in-vehicle interfaces.
- We evaluate our protocol in the high-fidelity PTV Vissim simulator on a real-world road network under diverse traffic conditions. Compared to baseline distributed control strategies, **our approach improves network throughput by 5%, reduces average vehicle delay by 13%, and lowers total stops by up to 5%,** while maintaining smooth and robust traffic flow.

Taken together, our results highlight a novel and practical traffic management paradigm: learning global speed modulation policies that achieve system-level coordination without infrastructure changes. By aligning with the capabilities of autonomous vehicles and real-time routing platforms, our method offers a feasible path to deployment through lightweight, app-level integration. Large-scale mobile telematics platforms like CMT (Cambridge Mobile Telematics 2025) and Samsara (Samsara Inc. 2025) are already adopted by tens of millions of drivers receiving instantaneous coordination and feedback from cloud-AI powered telematics systems, thereby underscoring the deployability of our approach and bridging ML innovation with scalable impact in modern traffic systems.

Related Work

The problem of road traffic management, especially freeway traffic management has been approached from many different angles across several research communities.

Traditional freeway management techniques aim to optimize flow, reduce congestion, and enhance safety through methods such as ramp metering (Arnold et al. 1998; Papageorgiou and Kotsialos 2002; Shaaban, Khan, and Hamila 2016), lane metering (Karimi Shahri et al. 2019; Hussain, Ghiasi, and Li 2016), and dynamic tolling (Mattsson 2008; Croci 2016; Yang, Purevjav, and Li 2020; May and Milne 2000). These infrastructure-heavy methods coordinate traffic through signaling, lane controls, and pricing strategies, but have proven inadequate in handling growing suburban congestion (Gordon et al. 1996). In contrast, more recent traffic management proposals fall into two broad categories: network-based approaches (for Advanced Transit and PATH; California Department of Transportation 2019; May et al. 2005; Texas A&M Transportation Institute (TTI) 2021), which modify physical infrastructure, and vehicle-based approaches like Uber and Google Maps (Nguyen 2015; Lau 2020), which provide real-time routing feedback. However, existing vehicle-based systems focus on individual optimization rather than system-level coordination. Our work addresses this gap by introducing a vehicle-based traffic control framework that learns global policies for network-wide regulation, without requiring new infrastructure.

Speed modulation as a method of road traffic management has been previously studied in (OH and Oh 2005; Soriguera, Torné, and Rosas 2013). OH and Oh (2005) use traditional theoretical models of free-flow traffic and derive the optimal speed control mechanism, but do not test the performance of their model with either empirical or simulation experiments. Soriguera, Torné, and Rosas (2013), on the other hand, assess the impact of a particular case study in Barcelona, and do not propose any optimal strategy. More recent efforts have applied reinforcement learning to variable speed limit control (Zhu and Ukkusuri 2014; Li et al. 2017), but these methods rely on macroscopic or link–queue formulations that simplify traffic flow and omit microscopic dynamics such as acceleration, merging, and car-following behavior. Furthermore, Zhu and Ukkusuri (2014) model only aggregated link dynamics, and Li et al. (2017) focus on localized bottleneck control, whereas our work achieves infrastructure-free, network-level coordination across multiple unsignalized highway segments using the PTV Vissim microscopic simulator, which captures individual driver behavior and realistic congestion formation.

Prior works have also tackled the problem of road traffic management using reinforcement learning (Prabuchandran, AN, and Bhatnagar 2014; Wiering et al. 2000; Mousavi, Schukat, and Howley 2017; Garg, Chli, and Vogiatzis 2018; Chu et al. 2019; Li et al. 2021; Chauhan, Bansal, and Sen 2020; Chauhan and Sen 2023, 2024), but most of the works in this domain apply reinforcement learning for learning optimal traffic signal controls, and do not apply to freeways. In the context of free-flow road networks, the applications of reinforcement learning have been studied mostly with cooperative vehicular networks (Pan et al. 2021; Peng et al. 2021; Shi et al. 2021; Vrbanić et al. 2021), which employ vehicle-level speed control and lane management strategies for traffic flow management. However, such approaches run into scalability issues since the multi-agent reinforcement learning model is dependent on peer-to-peer communication that scales super-linearly with the number of vehicles. Our system addresses this very important limitation, bringing our evaluation scale close to real-world deployment.

Pivoting to deployment, the human-in-the-loop deployment design of our system is inspired by prior work on advisory autonomy (Cho et al. 2023; Fridman 2018) and green light optimized speed advisory (GLOSA) systems (Stevanovic, Stevanovic, and Kergaye 2013). Adaptive signaling systems like SCATS and SCOOT (Stevanovic, Kergaye, and Martin 2009; Transport for New South Wales 2024; Transport Research Laboratory (TRL) 2024; Lowrie 1982; Hunt et al. 1981), which we initially planned to use as baselines for comparison, but are proprietary protocols that require external licenses that we were unable to procure, are smart city-scale traffic signaling systems that have seen limited deployment.

Preliminaries

In this section, we introduce concepts that play a key role in the design of our reinforcement learning system.

Traffic Flow Theory We focus on free-flow road networks, which lack traffic signals and operate under a maximum speed limit. For a single free-flow segment, traffic dynamics

follow the volume-density relationship $C = b \cdot v$, where C is flow volume, b is density, and v is velocity. Unlike traditional fluid models, traffic flow accounts for compressibility (variable density) and collision avoidance. The speed-density relationship is modeled via formulations by Greenshields et al. (1935), Underwood (1961), and Drake (1967), all of which yield a unimodal volume-density curve with a maximum at a critical density—validated empirically in Hall, Allen, and Gunter (1986). In multi-segment networks with merges and crossings, conflicting streams complicate the dynamics. These are typically resolved either via signals or through gap acceptance models (Lieu 1999), such as those by Troutbeck and Brilon (1996), McDonald and Armitage (1978), Tanner (1962), and Harders (1968) for unsignalized intersections. These models assume that drivers assess the safety of proceeding by identifying a time-based "gap" in traffic—measured as the interval between vehicles. While effective conceptually, such models rely on simplified geometric assumptions and become analytically intractable under complex real-world intersection conditions.

Traffic Simulation Framework Theoretical models of traffic flow are encoded into simulators for studying complex real-world traffic behavior. Macroscopic simulators (Papageorgiou et al. 2010; PTV Group 2023b) are generally used for studying aggregated variables such as density, flow, and speed at scale, using continuum traffic flow mechanics (Lighthill and Whitham 1955; Richards 1956; Payne 1973; Whitham 2011). Microscopic simulators, in contrast, simulate individual vehicles governed by rule-based car-following models (Hunter et al. 2021; Weidmann 1993; Newell 1961). This approach suits our RL-based setup, where traffic flow is modulated at the vehicle level. For our simulations, we use PTV Vissim (PTV Group 2023a), a state-of-the-art microscopic traffic simulator employing the Weidmann 99 driving model (Hunter et al. 2021), a psycho-physical behavior model where drivers adjust actions based on discrete thresholds of gap, speed, and relative speed. Previously, Fellendorf and Vortisch (2001) have shown that PTV Vissim can be calibrated across diverse real-world traffic scenarios to match both microscopic and macroscopic traffic patterns. We tuned the key Weidmann parameters to reproduce realistic speed-density-flow relationships as shown in in Figure 1 (top-left), and explained in detail in the Appendix. This approach ensures that the results are generalizable to other simulation frameworks and empirical settings.

Physics-Based Proof of Concept Traffic management can be framed as a network performance optimization problem targeting metrics like average travel time, throughput, and fairness, with a key objective of avoiding traffic jams (Lieu 1999). Prior work (Bhardwaj et al. 2023) has shown that even sub-critical increases in input flow can trigger persistent congestion once critical density is exceeded. Our strategy is thus based on two principles: (1) maintain density below critical thresholds, and (2) maximize flow speed.

To reconcile these, we apply a simplified backpressure-based speed modulation protocol inspired by queue-based routing algorithms from communication networks (Tassiulas and Ephremides 1990; Ribeiro 2009; Yoo et al. 2011; Seferoglu and Modiano 2015), and adapted to traffic contexts (Ma et al. 2020; Chang et al. 2020; Hu and Smith 2019). In a toy 2:1 merge scenario, let $C_\ell(t) = b_\ell(t) \cdot v_\ell(t)$ denote flow on segment ℓ. The density change on the downstream segment ℓ' is given by

$$\frac{db_{\ell'}}{dt} = \sum_{\ell_i} C_{\ell_i}(t) - C_{\ell'}(t),$$

and its remaining capacity before reaching critical density $B^*_{\ell'}$ is $\hat{b}_{\ell'}(t) = B^*_{\ell'} - b_{\ell'}(t)$. If δt is the control interval, and $\frac{db_{\ell'}}{dt} \cdot \delta t > \hat{b}_{\ell'}(t)$, input flows are scaled using:

$$\alpha = \frac{\hat{b}_{\ell'}(t)/\delta t + C_{\ell'}(t)}{\sum_{\ell_i} C_{\ell_i}(t)},$$

where α applies proportional backpressure via velocity modulation. Using PTV Vissim, we implemented this protocol and compared it to SCATS, SCOOT[1], equal-time, and no-control baselines on a single-lane merge topology. As shown in Figure 1, backpressure modulation outperformed baselines in throughput and time to jam formation, while intentionally reducing average speed.

However, performance varied with network geometry and driving dynamics. For instance, in single-lane merges, Weidmann 99 and gap acceptance behavior introduced a "gap penalty" at intersections, where vehicles slow down to assess safety before merging. Under high load, this leads to queuing and congestion. We also observed sensitivity to lane configurations, gap timing, and intersection geometry, highlighting the need for more adaptive, scalable strategies.

Methodology

Traffic control in free-flow road networks poses a complex and stochastic optimization challenge. Designing a deterministic speed-modulation policy using first-principles physics is intractable due to nonlinear dynamics, unmodeled driver behavior, and topology-induced variability. Instead, we adopt a reinforcement learning (RL) approach that enables learning from simulation, while embedding traffic-theoretic priors—such as volume-density relationships and gap acceptance theory—into the state and reward design. This hybridization allows for both interpretability and adaptability across varying traffic conditions. Our objective is to learn a control policy that dynamically adjusts segment-wise speed limits to maximize throughput and prevent congestion, without relying on fixed signaling infrastructure.

Environment Setup & Settings To reduce the modeling burden of creating complex traffic networks in PTV Vissim, we developed a semi-automated pipeline that converts real-world road layouts from OpenStreetMap (OpenStreetMap Foundation 2024) into Vissim-compatible formats. Using Overpass Turbo (Raifer 2024), we query and extract a filtered road network based on selected coordinates. The extracted data is then converted to OpenDrive format via `osm2xodr` (Meusener and username) 2022), which Vissim can import. While the conversion is not perfect, it significantly reduces the manual effort required to reconstruct large-scale road networks. After

[1]SCATS and SCOOT were simplified for single intersection and implemented based on original papers.

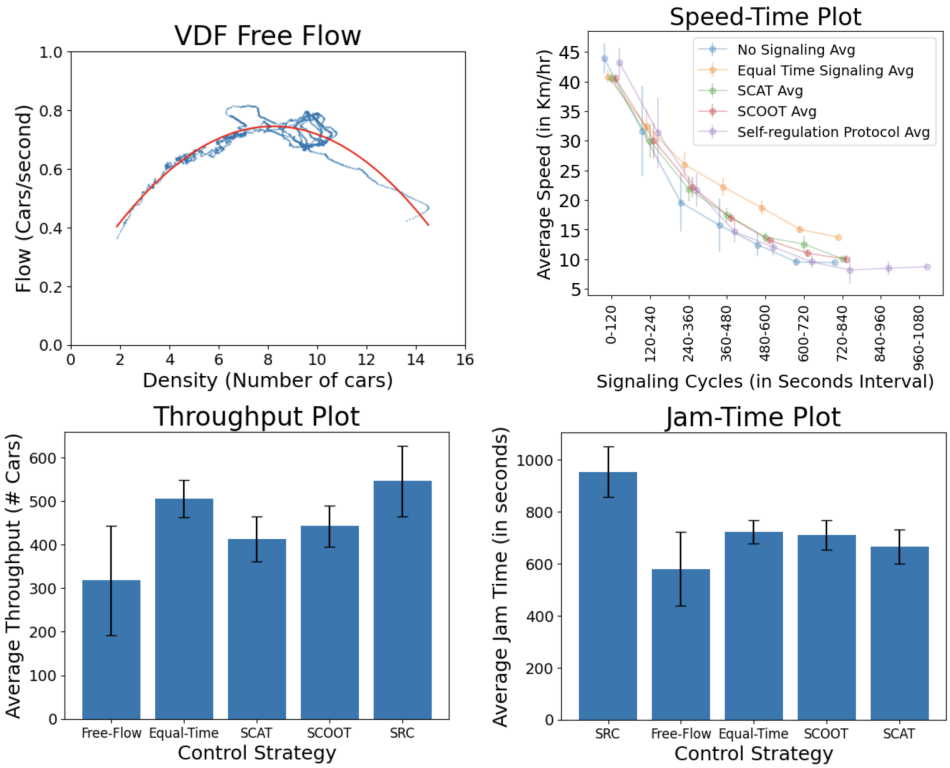


Figure 1: Flow-Density function plot (top-left) showing the instantaneous flow rate as a function of the queue size. The plot was experimentally derived by simulating a single segment in PTV Vissim. Blue dots are the observed points and the red line is the fitted plot. Other plots show the performance of different traffic control strategies on single lane merge topology on average speed, throughput, and time to hit a jam.

generating the network layout, we partition the road into continuous super-segments, each spanning 2–3 kilometers, to serve as control units that capture free-flow traffic characteristics such as volume-density relationships (Nishinari 2014; Kerner 2004). OpenStreetMap (OSM) segments are typically much shorter, reflecting changes in road attributes, intersections, or administrative boundaries, and are optimized for geographic accuracy rather than traffic analysis. To construct super-segments, we manually aggregated contiguous OSM links using sampled vehicle trajectories from PTV Vissim simulations (west $\rightarrow$ east). Intersections, entry/exit nodes, or natural boundaries were used as division points, and aggregation was stopped once the segment length exceeded roughly 3–5 km. This procedure ensured that each super-segment represented a coherent roadway section suitable for stable and interpretable control. The RL model operates at the super-segment level, determining speed modulation policies that are fed back to the simulator. Segment-level states within each super-segment are averaged to form a composite state vector, which serves as the model input and enables efficient, interpretable control.

Reinforcement Learning Model Architecture We formulate traffic flow control as a Markov Decision Process (MDP), where a centralized agent observes the global traffic state and modulates the maximum allowable velocity on each of N super-segments to optimize throughput and avoid congestion. The MDP is defined by the tuple $\mathcal{M} = (\mathcal{S}, \mathcal{O}, \mathcal{A}, \mathcal{P}, r, \pi, \gamma)$ as follows:

State and Observation ($\mathcal{S}, \mathcal{O}$): At time t, each super-segment $i \in \{1, \ldots, N\}$ is described by a local state vector $s_i^t \in \mathbb{R}^5$ containing key traffic metrics: vehicle density ρ_i^t, average speed v_i^t, average gap g_i^t, inflow rate λ_i^t, and outflow rate μ_i^t. These features provide a localized snapshot of congestion, stability, and flow pressure. The global observation is the concatenation $o^t = [s_1^t, \ldots, s_N^t] \in \mathbb{R}^{5N}$, forming the input to the agent.

Action Space ($\mathcal{A}$): The agent selects an action vector $a^t = [a_1^t, \ldots, a_N^t]$, where $a_i^t \in \mathcal{A}_i$ represents maximum allowable speed (in km/h) for super-segment i during $[t, t + \Delta t]$. We use a fixed, discrete action set $\mathcal{A}_i$ across all segments to promote interpretability, robustness, and compatibility with human-in-the-loop implementations such as in-vehicle advisories.

Transition Dynamics ($\mathcal{P}$): The traffic state evolves stochastically according to $S^{t+1} \sim \mathcal{P}(S^{t+1} \mid S^t, A^t)$, where vehicle interactions are governed by the Weidmann 99 car-following model implemented in the PTV Vissim simulator.

Reward Function (r): The reward r_i^t for each super-segment i at time t combines two objectives. First, a penalty is applied if the density exceeds a critical threshold $\rho^* = 0.3$:

$$r_i^{(1)}(t) = \begin{cases} -\alpha_d & \text{if } \rho_i^t > \rho^*, \\ 0 & \text{otherwise,} \end{cases}$$

where α_d is a large penalty constant. Second, the agent is rewarded for maintaining high average speed:

$$r_i^{(2)}(t) = \beta_v \cdot v_i^t,$$

where β_v is a positive scaling factor. The total reward for

segment i is $r_i^t = r_i^{(1)}(t) + r_i^{(2)}(t)$, and the global reward is $r^t = \sum_{i=1}^{N} r_i^t$. This design simulates a backpressure-like mechanism, where high-density regions act as implicit pressure signals and speed rewards encourage smooth flow.

Objective and Policy (π, γ): The agent aims to learn a policy $\pi : \mathcal{O} \to \mathcal{A}$ that maximizes the expected cumulative discounted return:

$$\pi^* = \arg\max_{\pi} \mathbb{E}_{\pi} \left[\sum_{t=0}^{T} \gamma^t r^t \right],$$

where $\gamma \in [0, 1]$ is the discount factor.

To learn the optimal policy, we employ Deep Q-Learning (DQN) (Mnih et al. 2015). This value-based method is well suited to our centralized setting and discrete action space, offering faster convergence and higher sample efficiency, compared to actor-critic or policy-gradient methods often used in continuous or multi-agent scenarios (Tan 2025). The Q-function $Q(o^t, a^t; \theta_n)$ is parameterized by a neural network with weights θ_n, and trained by minimizing the temporal-difference loss:

$$\mathcal{L}(\theta_n) = \mathbb{E}_{(o^t, a^t, r^t, o^{t'})} \left[\sum_{i=1}^{N} \left(r_i^t + \gamma \max_{a'} Q(o^{t'}, a_i'; \theta_{n-1}) \quad - Q(o^t, a_i^t; \theta_n) \right)^2 \right],$$

where transitions $(o^t, a^t, r^t, o^{t'})$ are sampled from a replay buffer. This formulation supports interpretable learning of control policies that balance throughput and congestion mitigation across dynamic traffic conditions.

Throughput-Optimality under Two-Regime Flow Dynamics Let λ_i^t, μ_i^t denote the inflow and outflow in super-segment i at time t, and let $f(\rho_i^t)$ be the empirically observed flow-density function. Assume that (i) $f(\rho)$ is unimodal with a unique maximum at critical density ρ^*; (ii) The environment evolves as $\rho_i^{t+1} = \rho_i^t + \lambda_i^t - \mu_i^t$; (iii) There is sufficient and persistent inflow: $\lambda_i^t \geq \lambda_{\min} \forall t$; and $\lambda_{\min}$ is large enough that the system would reach ρ^* if outflow were not properly managed. Then, the optimal policy π^* that maximizes the expected discounted return, is throughput-optimal in the sense in the sense that it maintains the network in a stable equilibrium near ρ^*, the point of maximum sustainable flow, while preventing congestion collapse.

To justify this claim, we divide the analysis into two regimes based on the density ρ_i^t in each segment.

Free-flow regime ($\rho_i^t < \rho^*$): In this region, the traffic dynamics are dominated by car-following behavior. The vehicle flow is given by

$$\mu_i^t = \frac{v_i^t}{d_0 + v_i^t T}$$

where d_0 is the standstill distance and T is the headway time. This function is strictly increasing in v_i^t for $\rho_i^t < \rho^*$ since

$$\frac{d\mu_i^t}{dv_i^t} = \frac{d_0}{(d_0 + v_i^t T)^2} > 0$$

Hence, in the free-flow regime flow increases with velocity. Now consider the agent's reward:

$$r_i^t = \beta_v \cdot v_i^t \quad \text{(since } \rho_i^t < \rho^* \Rightarrow \text{no penalty)}$$

So maximizing r_i^t in this region leads to maximizing μ_i^t.

Congested regime ($\rho_i^t > \rho^*$): In this region, emergent traffic behaviors dominate (e.g., braking waves), and empirical VDFs (e.g., Figure 1(top-left)) show that: $\frac{d\mu_i^t}{d\rho_i^t} < 0$

Additionally, the reward includes a penalty term:

$$r_i^t = -\alpha_d + \beta_v \cdot v_i^t$$

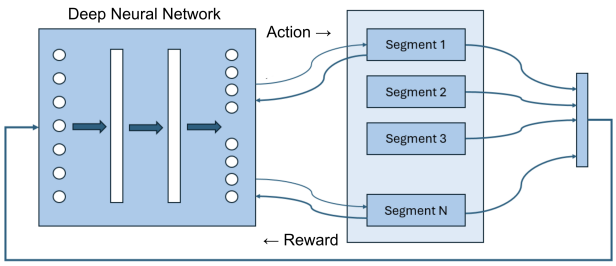
The large value of α_d and bounded velocity in the action set ensures that r_i^t is strictly negative. Hence, the agent has incentive to reduce ρ_i^t to exit this regime, but this behavior also maximizes the outflow.

Simulation-Based Training and Transfer Learning We train a deep Q-learning model to learn speed modulation policies for self-regulating cars using the PTV Vissim microscopic simulator. The agent controls 12 super-segments, each with three discrete speed options, and is implemented as a two-layer neural network. Every minute, the agent observes a state vector for each super-segment containing density, average speed, gap, inflow, and outflow, and selects speed actions accordingly. Training is performed over simulation episodes of either 1200 or 2500 seconds, using an initial exploration rate of 0.9 that decays to 0.1 (with decay rate 0.02–0.05 per episode), and a discount factor $\gamma = 0.9$ to encourage long-term optimization.

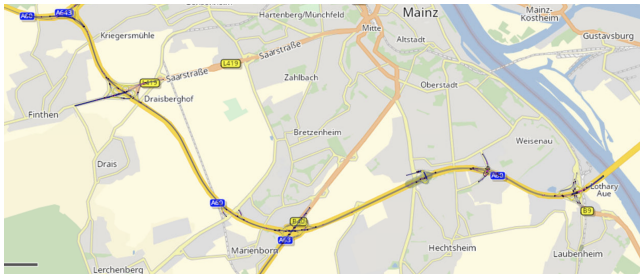
We evaluate two regimes. In the Self-Regulating Cars (SRC) protocol, the model is trained directly on a high-traffic input pattern over 2500 seconds with the action set of (30, 45, or 60 km/h). In the transfer learning variant (SRC-TL), the agent is first trained on a low-traffic input pattern for 1200 seconds to promote exploration across traffic states with the action set of (20, 40, or 60 km/h), then transferred—without fine-tuning—to a different high-traffic pattern for evaluation over 2500 seconds. This setup improves policy robustness by enabling convergence under stable dynamics while ensuring generalization across flow patterns. The use of coarsely discretized actions aids deployability via human-in-the-loop implementations (e.g., mobile advisory systems), and the model is trained solely on instantaneous state and reward signals. This design deliberately avoids temporal input features to ensure transferability across varying demand scenarios, leveraging the consistency of underlying traffic physics such as flow-density and gap acceptance relationships.

Evaluation

Evaluation Testbed In free-flow road networks, traffic jams rarely occur on uninterrupted straight segments unless capacity changes. While phantom jams (Flynn et al. 2009) can arise, they are unlikely in our setting where self-driving agents adhere strictly to the Weidmann 99 model. Thus, we focus on applying speed modulation only at network bottlenecks—typically during morning peak hours when suburban traffic converges at city entry points, causing congestion. To study such conditions, we modeled a real-world highway segment in Mainz, Germany (Figure 2b), simulating eastbound rush hour traffic with multiple ramp inflows. The input pattern was designed to reflect realistic cumulative west-to-east flow, where upstream vehicles continue forward and additional traffic merges at successive ramps. The



(a) RL-based model to control velocity on different road segments.



(b) Real-world free-flow road network layout in Mainz, Germany.

Figure 2: RL framework and deployment environment.

A643 & A60 junction introduced 3,000–5,000 veh/hr, representing long-distance freeway traffic. L419 (Draisberghof, Saarstraße) added 1,500–3,000 veh/hr from arterial roads. Further east, the A60 & A63 Marienborn interchange contributed 4,000–6,000 veh/hr from a major highway junction. Finally, the A60 & Hechtsheim ramp merged 2,500–4,000 veh/hr from urban arterials. No new vehicles entered beyond this point, and all traffic exited via the B9 & A60 Mainz bridge. This setup allowed Vissim to simulate realistic merging dynamics, congestion buildup & rising density near urban core without artificially overloading any single entry point.

Baselines Our goal is to evaluate a centralized speed-modulation controller that operates without introducing any new roadside infrastructure. The system assumes that each vehicle has a mobile phone or in-vehicle device with internet connectivity, allowing a central service to broadcast recommended speeds in real time. While the idea of speed modulation has been explored conceptually in prior works (OH and Oh 2005; Soriguera, Torné, and Rosas 2013), these efforts have largely focused on localized or analytical studies rather than deployable solutions evaluated on realistic road networks. As a result, there are no existing large-scale, infrastructure-free speed-modulation baselines to compare against. In contrast, adaptive metering and decentralized RL strategies (Chen et al. 2020; Wei et al. 2019) typically rely on additional infrastructure such as ramp meters, lane-specific actuators, roadside sensors, or vehicle-to-infrastructure communication. Including such methods as baselines would implicitly assume hardware upgrades that our framework explicitly seeks to avoid. So, we compare against the infrastructure-based control paradigms that already dominate real-world deployments, using only existing signal infrastructure.

We consider four such baselines: `no control`, `equal split`, `green wave`, and `proportional split` signaling. The `no control` baseline represents the simulator's default behavior. In `equal split` signaling, each incoming road receives an equal share of a 120-second signal cycle with a 5-second intergreen interval. The `green wave` configuration introduces offsets between consecutive intersections so that vehicles traveling at the speed limit encounter mostly green lights. In `proportional split` signaling, green times are allocated based on measured traffic loads, allowing busier intersections proportionally longer green durations. These baselines were implemented by embedding traffic signals into the network layout; they modulate only how many vehicles can enter the main highway and do not directly control vehicle speeds. This setup allows us to highlight the operational and social advantages of a connectivity-based, software-layer controller over traditional infrastructure-centric systems.

Metrics We evaluate traffic performance using 14 metrics from PTV Vissim, which we organize into five categories: throughput, flow smoothness, delay, congestion, & emissions.

Throughput is captured by `VEHARR` (vehicles that completed trips), `DISTTOT` (total distance traveled), `SPEEDAVG` (average traffic flow speed) and `DEMANDLATENT` (unfulfilled demand).

Flow smoothness is measured by `STOPSAVG` and `STOPSTOT`, representing the average and total number of stops per vehicle, respectively.

Delay-related metrics include `DELAYAVG`, `DELAYSTOPAVG`, `DELAYTOT`, `DELAYSTOPTOT`, and `DELAYLATENT`. While we report these for completeness, they are influenced by simulation implementation details and should be interpreted cautiously.

Congestion is quantified via `VEHACT`, the number of active vehicles remaining in the network, and `TRAVTMAVG`, which we obtain by normalizing total recorded travel time in vissim (`TRAVTMTOT`) by total vehicle count (`VEHARR + VEHACT`). A lower value here reflects less congestion.

Emissions are represented by `EMISSIONSCO2`, estimated from vehicle speed using the method of Ntziachristos and Samaras (2000). The total calculated emissions are further normalized by multiplying the recorded emissions with (`VEHARR + VEHACT + DEMANDLATENT`/`VEHARR + VEHACT`) to account for the latent demand.

Taken together, these metrics provide a holistic view of performance of a traffic network.

Results Table 1 provides the performance results of our protocol and the baselines on the aforementioned metrics and testbed. Our self regulating cars protocol (SRC) and transfer learned (SRC(TL)) with a different training input pattern, achieve the best throughput performance across all four metrics (`VEHARR, DISTTOT, SPEEDAVG, DEMANDLATENT`), aligning with our primary design objective of maximizing network capacity and preventing traffic jams. Our protocol also achieves greatest smoothness in the traffic flow by minimizing the stop frequency (`STOPSAVG, STOPSTOT`). We also note competitive delay and congestion performance achieved by our protocol with best performance achieved in `DELAYAVG, DELAYTOT`, and `TRAVTMAVG`. While Proportional Split baseline achieves lowest active vehicle numbers (`VEHACT`), we observed traffic jams and high congestion on entry points to the main highway. Our protocol achieves the best `VEHACT` with a more

Metric	No Control	Equal Split	Proportional Split	Green Wave	SRC	SRC (TL)
Throughput						
↑ VEHARR (#)	2227 ± 19	2002 ± 22	1873 ± 10	2045 ± 17	**2337.8 ± 16.42**	2291.6 ± 33.78
↑ DISTTOT (10^3 km)	12.300 ± 0.315	11.723 ± 0.086	9.598 ± 0.083	11.904 ± 0.050	**13.078 ± 0.099**	12.467 ± 0.065
↓ DEMANDLATENT (#)	$1,722.8 \pm 43.73$	$1,746.8 \pm 27.53$	$2,164.8 \pm 32.51$	$1{,}570.8 \pm 20.36$	**1,547.8 ± 35.074**	$1,643.8 \pm 57.45$
↑ SPEEDAVG (km/hr)	13.16 ± 0.42	11.22 ± 0.05	10.39 ± 0.07	10.70 ± 0.10	**13.23 ± 0.16**	13.06 ± 0.06
Flow Smoothness						
↓ STOPSAVG (#)	15.17 ± 0.34	32.31 ± 0.50	23.81 ± 0.52	199.76 ± 3.19	15.94 ± 0.96	**14.51±0.32**
↓ STOPSTOT (10^3 #)	65.425 ± 1.711	139.175 ± 2.059	92.151 ± 2.345	895.682 ± 19.05	71.552 ± 4.165	**63.723 ± 1.40**
Delay						
↓ DELAYAVG (sec)	571.42 ± 6.50	689.11 ± 3.39	683.65 ± 9.20	714.00 ± 6.95	527.20 ± 9.46	**499.36 ± 7.96**
↓ DELAYSTOPAVG (sec)	**226.26 ± 7.31**	460.62 ± 3.79	473.28 ± 9.88	395.73 ± 5.39	306.05 ± 11.33	251.93 ± 12.38
↓ DELAYTOT (10^6 sec)	2.464 ± 0.039	2.968 ± 0.021	2.646 ± 0.042	3.201 ± 0.049	2.366 ± 0.039	**2.193 ± 0.030**
↓ DELAYSTOPTOT (10^6 sec)	**0.976 ± 0.034**	1.984 ± 0.025	1.832 ± 0.044	1.774 ± 0.030	1.373 ± 0.051	1.106 ± 0.054
↓ DELAYLATENT (10^6 sec)	5.618 ± 0.107	5.467 ± 0.076	6.023 ± 0.088	**5.132 ± 0.069**	5.334 ± 0.097	5.511 ± 0.117
Congestion						
↓ VEHACT (#)	2085 ± 38	2305 ± 32	**1997 ± 11**	2438 ± 55	2150 ± 27.23	2100.2 ± 14.92
↓ TRAVTMAVG (10^3 sec/veh)	**0.780 ± 0.007**	0.873 ± 0.006	0.859 ± 0.011	0.894 ± 0.012	0.793 ± 0.009	**0.783 ± 0.006**
Emissions						
↓ EMISSIONSCO2 (g/m, normalized)	0.525 ± 0.013	0.600 ± 0.001	**0.513 ± 0.005**	3.001 ± 0.023	0.530 ± 0.005	0.526 ± 0.003

Table 1: Vehicle network performance measurements for self-regulating cars and baseline traffic management strategies for real-world representative traffic loads. The variance and mean values were calculated over 5 random seeds, with each experiment running for 2500 seconds. Arrow directions show direction of desired outcome. Our protocol dynamically adjusts desired speeds, unlike fixed-speed baselines. Because Vissim computes delay relative to a speed-dependent free-flow travel time, raw delay values are not directly comparable. We therefore evaluate delay at equivalent congestion levels and rely on additional network-wide efficiency metrics for fairness.

Metric	75% control	50% control	25% control
Throughput			
↑ VEHARR (10^3)	2.27 ± 0.02	2.25 ± 0.06	2.22 ± 0.04
↑ DISTTOT (10^4 km)	1.26 ± 0.02	1.23 ± 0.02	1.20 ± 0.01
↓ DEMANDLATENT (10^3)	1.65 ± 0.06	1.72 ± 0.05	1.69 ± 0.05
↑ SPEEDAVG (10 km/hr)	1.35 ± 0.01	1.32 ± 0.01	1.29 ± 0.00
Flow Smoothness			
↓ STOPSAVG (10)	1.45 ± 0.07	1.46 ± 0.07	1.42 ± 0.04
↓ STOPSTOT (10^5)	0.63 ± 0.03	0.63 ± 0.02	0.62 ± 0.01
Delay			
↓ DELAYAVG (10^2 s)	5.40 ± 0.07	5.53 ± 0.14	5.62 ± 0.08
↓ DELAYSTOPAVG (10^2 s)	2.56 ± 0.11	2.42 ± 0.11	2.40 ± 0.10
↓ DELAYTOT (10^6 s)	2.35 ± 0.02	2.38 ± 0.05	2.44 ± 0.02
↓ DELAYSTOPTOT (10^6 s)	1.11 ± 0.05	1.04 ± 0.05	1.04 ± 0.04
↓ DELAYLATENT (10^6 s)	5.53 ± 0.11	5.64 ± 0.07	5.65 ± 0.09
Congestion			
↓ VEHACT (10^3 veh)	2.08 ± 0.03	2.06 ± 0.04	2.12 ± 0.04
↓ TRAVTMAVG (10^3 s/veh)	0.78 ± 0.01	0.78 ± 0.01	0.77 ± 0.01

Table 2: Ablation experiments with mixed traffic inputs. Results show a clear trend of improvement in throughput and flow metrics with increasing fraction of compliant vehicles.

uniformly distributed vehicle profile. The `EMISSIONSC02` numbers are calculated under lots of approximation assumptions, and the associated variance in calculations is higher than the difference between best emissions and our numbers. Consistent with earlier expectations, our experimental results show that the self-regulating cars protocol performs well even when evaluated on an unseen traffic input pattern. The transferred policy (SRC-TL) continues to optimize key performance metrics, such as throughput and delay, and maintains competitive results across other indicators when compared to both learned and baseline signaling strategies.

Mixed Traffic Ablations We conducted ablation experiments on the same network layout, where only a subset of vehicles is controlled by our RL model (25%, 50%, and 75%). All other configuration parameters remained unchanged; however, introducing mixed autonomy led to minor variations in simulation behavior. The results shown in Table 2 demonstrate a consistent trend: as the proportion of controlled vehicles increases, network performance improves significantly. Specifically, average speed and throughput increase, while overall delay and congestion-related metrics decrease. These findings reinforce the practical applicability of our method in settings with partial adoption. Similar trends are observed in stochastic conformity settings.

Conclusion, Limitations, and Future Work

We introduced a novel traffic control paradigm based on self-regulating cars, smart vehicles that coordinate to maintain smooth traffic flow in free-flow networks by adjusting speeds according to a physics-informed reinforcement learning policy. Our method achieves comparable or superior performance to existing signaling protocols in high-fidelity real-world simulations, and can be deployed via lightweight integration with in-vehicle navigation systems. A key current limitation is the manual specification of road networks for simulation, restricting our evaluation to a single topology. Moreover, the protocol's effectiveness relies on a sufficient adoption rate of compliant vehicles, and its current design allows non-compliant agents to gain a strategic advantage. In future work, we aim to automate the simulator input pipeline, develop a city scale simulation engine, and extend the control framework to incorporate lane-level decision-making to design mechanisms to mitigate adversarial behavior.

Ethics Statement

This study is conceptual and simulation-based, and does not involve deployment or interaction with real vehicles or human subjects. Any real-world implementation of self-regulating traffic control systems would need to be conducted in collaboration with governmental or regulatory agencies to ensure transparency, fairness, and public accountability. Care must be taken to prevent disproportionate impacts across different vehicle types, communities, or socioeconomic groups, and to ensure equitable access to the benefits of connected and autonomous vehicle technologies. These aspects represent important directions for future deployment-oriented work.

Acknowledgements

Ankit Bhardwaj and Lakshminarayanan Subramanian were supported by the NSF Grant (Award Number 2335773) titled "EAGER: Scalable Climate Modeling using Message-Passing Recurrent Neural Networks". Lakshminarayanan Subramanian was also funded in part by the NSF Grant (award number OAC-2004572) titled "A Data-informed Framework for the Representation of Sub-grid Scale Gravity Waves to Improve Climate Prediction".

References

Arnold, E.; et al. 1998. Ramp metering: a review of the literature.

Badger, E. 2013. How traffic congestion affects economic growth. *Recuperado el*, 18.

Bhardwaj, A.; Iyer, S. R.; Ramesh, S.; White, J.; and Subramanian, L. 2023. Understanding sudden traffic jams: From emergence to impact. *Development Engineering*, 8: 100105.

California Department of Transportation. 2019. PeMS: California Performance Measurement System. http://pems.dot.ca.gov/. Accessed: 2024-10-21.

Cambridge Mobile Telematics. 2025. Cambridge Mobile Telematics. Accessed: 2025-05-15.

Chang, W.; Roy, D.; Zhao, S.; Annaswamy, A.; and Chakraborty, S. 2020. CPS-oriented modeling and control of traffic signals using adaptive back pressure. In *2020 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 1686–1691. IEEE.

Chauhan, S.; Bansal, K.; and Sen, R. 2020. EcoLight: Intersection control in developing regions under extreme budget and network constraints. *Advances in Neural Information Processing Systems*, 33: 13027–13037.

Chauhan, S.; and Sen, R. 2023. REALLIGHT: DRL based Intersection Control in Developing Countries without Traffic Simulators. In *NeurIPS 2023 Computational Sustainability: Promises and Pitfalls from Theory to Deployment*.

Chauhan, S. K.; and Sen, R. 2024. FrugalLight: Symmetry-Aware Cyclic Heterogeneous Intersection Control using Deep Reinforcement Learning with Model Compression, Distillation and Domain Knowledge. *ACM Journal on Computing and Sustainable Societies*, 2(2): 1–32.

Chen, C.; Wei, H.; Xu, N.; Zheng, G.; Yang, M.; Xiong, Y.; Xu, K.; and Li, Z. 2020. Toward a thousand lights: Decentralized deep reinforcement learning for large-scale traffic signal control. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, 3414–3421.

Cho, J.-H.; Li, S.; Kim, J.; and Wu, C. 2023. Temporal transfer learning for traffic optimization with coarse-grained advisory autonomy. *arXiv preprint arXiv:2312.09436*.

Chu, T.; Wang, J.; Codecà, L.; and Li, Z. 2019. Multi-agent deep reinforcement learning for large-scale traffic signal control. *IEEE transactions on intelligent transportation systems*, 21(3): 1086–1095.

Croci, E. 2016. Urban road pricing: a comparative study on the experiences of London, Stockholm and Milan. *Transportation Research Procedia*, 14: 253–262.

Downie, A. 2008. The World's Worst Traffic Jams. http://content.time.com/time/world/article/0,8599,1733872,00.html.

Drake, J. S. 1967. A statistical analysis of speed density hypothesis. *HRR*, 154: 53–87.

Fellendorf, M.; and Vortisch, P. 2001. Validation of the microscopic traffic flow model VISSIM in different real-world situations. In *transportation research board 80th annual meeting*, volume 11, 7–11.

Flynn, M. R.; Kasimov, A. R.; Nave, J.-C.; Rosales, E.; and Seibold, B. 2009. Traffic modeling-Phantom traffic jams and traveling jamitons. *Traffic*, 8(29): 2016.

for Advanced Transit, C. P.; and (PATH), H. 1993. A Program of Research in Advanced Vehicle Control and Traffic Management for Intelligent Vehicle-Highway Systems (IVHS). Technical report, Institute of Transportation Studies, University of California, Berkeley, Berkeley, CA.

Fridman, L. 2018. Human-centered autonomous vehicle systems: Principles of effective shared autonomy. *arXiv preprint arXiv:1810.01835*.

Garg, D.; Chli, M.; and Vogiatzis, G. 2018. Deep reinforcement learning for autonomous traffic light control. In *2018 3rd IEEE international conference on intelligent transportation engineering (ICITE)*, 214–218. IEEE.

Gordon, R. L.; Reiss, R. A.; Haenel, H.; Case, E.; French, R. L.; Mohaddes, A.; Wolcott, R.; et al. 1996. Traffic control systems handbook. Technical report, United States. Federal Highway Administration. Office of Technology Applications.

Greenshields, B. D.; Bibbins, J. R.; Channing, W.; and Miller, H. H. 1935. A study of traffic capacity. In *Highway research board proceedings*, volume 14, 448–477. Washington, DC.

Hall, F. L.; Allen, B. L.; and Gunter, M. A. 1986. Empirical analysis of freeway flow-density relationships. *Transportation Research Part A: General*, 20(3): 197–210.

Harders, J. 1968. The capacity of unsignalized urban intersections. *Schriftenreihe Strassenbau und Strassenverkehrstechnik*, 76: 1968.

Hu, H.-C.; and Smith, S. F. 2019. Softpressure: A schedule-driven backpressure algorithm for coping with network congestion. *arXiv preprint arXiv:1903.02589*.

Hunt, P. B.; Robertson, D. I.; Bretherton, R. D.; and Winton, R. I. 1981. SCOOT-a Traffic Responsive Method of Coordinating Signals.

Hunter, M.; et al. 2021. VISSIM Simulation Guidance. Technical report, Georgia. Department of Transportation. Office of Performance-Based ....

Hussain, O.; Ghiasi, A.; and Li, X. 2016. Freeway lane management approach in mixed traffic environment with connected autonomous vehicles. *arXiv preprint arXiv:1609.02946*.

Karimi Shahri, P.; Chintamani Shindgikar, S.; HomChaudhuri, B.; and Ghasemi, A. H. 2019. Optimal lane management in heterogeneous traffic network. In *Dynamic Systems and Control Conference*, volume 59162, V003T18A003. American Society of Mechanical Engineers.

Kerner, B. S. 2004. Three-phase traffic theory and highway capacity. *Physica A: Statistical Mechanics and its Applications*, 333: 379–440.

Lau, J. 2020. Google Maps 101: How AI helps predict traffic and determine routes. https://blog.google/products/maps/google-maps-101-how-ai-helps-predict-traffic-and-determine-routes/.

Li, Z.; Liu, P.; Xu, C.; Duan, H.; and Wang, W. 2017. Reinforcement learning-based variable speed limit control strategy to reduce traffic congestion at freeway recurrent bottlenecks. *IEEE transactions on intelligent transportation systems*, 18(11): 3204–3217.

Li, Z.; Yu, H.; Zhang, G.; Dong, S.; and Xu, C.-Z. 2021. Network-wide traffic signal control optimization using a multi-agent deep reinforcement learning. *Transportation Research Part C: Emerging Technologies*, 125: 103059.

Lieu, H. 1999. Traffic-Flow Theory. volume 62 of *Public Roads*. Federal Highway Administration.

Lighthill, M. J.; and Whitham, G. B. 1955. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the royal society of london. series a. mathematical and physical sciences*, 229(1178): 317–345.

Lowrie, P. 1982. The Sydney coordinated adaptive traffic system - principles, methodology, algorithms.

Ma, D.; Xiao, J.; Song, X.; Ma, X.; and Jin, S. 2020. A back-pressure-based model with fixed phase sequences for traffic signal optimization under oversaturated networks. *IEEE Transactions on Intelligent Transportation Systems*, 22(9): 5577–5588.

Mattsson, L.-G. 2008. Road pricing: Consequences for traffic, congestion and location. In *Road pricing, the economy and the environment*, 29–48. Springer.

May, A. D.; Cayford, R.; Merritt, G.; and Leiman, L. 2005. The Berkeley Highway Laboratory. Path research report, California PATH, University of California, Berkeley.

May, A. D.; and Milne, D. S. 2000. Effects of alternative road pricing systems on network performance. *Transportation Research Part A: Policy and Practice*, 34(6): 407–436.

McDonald, M.; and Armitage, D. 1978. The capacity of roundabouts. *Traffic Engineering & Control*, 19(N10).

Meusener, J.-H.; and username), F. G. 2022. osm2xodr: OpenStreetMap to OpenDRIVE Converter. https://github.com/JHMeusener/osm2xodr. Accessed: October 21, 2024.

Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A. A.; Veness, J.; Bellemare, M. G.; Graves, A.; Riedmiller, M.; Fidjeland, A. K.; Ostrovski, G.; et al. 2015. Human-level control through deep reinforcement learning. *nature*, 518(7540): 529–533.

Mousavi, S. S.; Schukat, M.; and Howley, E. 2017. Traffic light control using deep policy-gradient and value-function-based reinforcement learning. *IET Intelligent Transport Systems*, 11(7): 417–423.

Newell, G. F. 1961. Nonlinear effects in the dynamics of car following. *Operations research*, 9(2): 209–229.

Nguyen, T. 2015. ETA Phone Home: How Uber Engineers an Efficient Route. https://www.uber.com/blog/engineering-routing-engine/.

Nishinari, K. 2014. Traffic flow dynamics: Data, models and simulation. *Physics Today*, 67(3): 54–54.

Ntziachristos, L.; and Samaras, Z. 2000. Speed-dependent representative emission factors for catalyst passenger cars and influencing parameters. *Atmospheric environment*, 34(27): 4611–4619.

OH, J.-S.; and Oh, C. 2005. Dynamic speed control strategy for freeway traffic congestion management. *Journal of the Eastern Asia Society for Transportation Studies*, 6: 595–607.

OpenStreetMap Foundation. 2024. OpenStreetMap: The Free Wiki World Map. https://www.openstreetmap.org. Data available under the Open Database License (ODbL). Accessed: October 21, 2024.

Pan, T.; Guo, R.; Lam, W. H.; Zhong, R.; Wang, W.; and He, B. 2021. Integrated optimal control strategies for freeway traffic mixed with connected automated vehicles: A model-based reinforcement learning approach. *Transportation research part C: emerging technologies*, 123: 102987.

Papageorgiou, M.; and Kotsialos, A. 2002. Freeway ramp metering: An overview. *IEEE transactions on intelligent transportation systems*, 3(4): 271–281.

Papageorgiou, M.; Papamichail, I.; Messmer, A.; and Wang, Y. 2010. Traffic simulation with METANET. *Fundamentals of traffic simulation*, 399–430.

Payne, H. J. 1973. Freeway traffic control and surveillance model. *Transportation Engineering Journal of ASCE*, 99(4): 767–783.

Peng, B.; Keskin, M. F.; Kulcsár, B.; and Wymeersch, H. 2021. Connected autonomous vehicles for improving mixed traffic efficiency in unsignalized intersections with deep reinforcement learning. *Communications in Transportation Research*, 1: 100017.

Prabuchandran, K.; AN, H. K.; and Bhatnagar, S. 2014. Multi-agent reinforcement learning for traffic signal control. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2529–2534. IEEE.

PTV Group. 2023a. PTV Vissim: Microscopic Traffic Flow Simulation Software. https://www.ptvgroup.com. Karlsruhe, Germany. Accessed 2024-10-21.

PTV Group. 2023b. PTV VISUM: Traffic Planning Software. https://www.ptvgroup.com. Karlsruhe, Germany. Accessed 2024-10-21.

Raifer, M. 2024. Overpass Turbo: A Web-Based Data Mining Tool for OpenStreetMap. https://overpass-turbo.eu/. Accessed: 2024-10-21.

Ribeiro, A. 2009. Stochastic soft backpressure algorithms for routing and scheduling in wireless ad-hoc networks. In *2009 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 137–140. IEEE.

Richards, P. I. 1956. Shock waves on the highway. *Operations research*, 4(1): 42–51.

Samsara Inc. 2025. Samsara: The Connected Operations Cloud. https://www.samsara.com/. Accessed: 2025-05-15.

Seferoglu, H.; and Modiano, E. 2015. Separation of routing and scheduling in backpressure-based wireless networks. *IEEE/ACM Transactions on Networking*, 24(3): 1787–1800.

Shaaban, K.; Khan, M. A.; and Hamila, R. 2016. Literature review of advancements in adaptive ramp metering. *Procedia Computer Science*, 83: 203–211.

Shi, H.; Zhou, Y.; Wu, K.; Wang, X.; Lin, Y.; and Ran, B. 2021. Connected automated vehicle cooperative control with a deep reinforcement learning approach in a mixed traffic environment. *Transportation Research Part C: Emerging Technologies*, 133: 103421.

Soriguera, F.; Torné, J. M.; and Rosas, D. 2013. Assessment of dynamic speed limit management on metropolitan freeways. *Journal of Intelligent Transportation Systems*, 17(1): 78–90.

Stevanovic, A.; Kergaye, C.; and Martin, P. 2009. SCOOT and SCATS: A Closer Look into Their Operations.

Stevanovic, A.; Stevanovic, J.; and Kergaye, C. 2013. Green Light Optimized Speed Advisory Systems: Impact of Signal Phasing Information Accuracy. *Transportation Research Record*, 2390(1): 53–59.

Tan, C. 2025. Comparative Study of Reinforcement Learning Performance Based on PPO and DQN Algorithms. *Applied and Computational Engineering*, 175: 30–36.

Tanner, J. 1962. A theoretical analysis of delays at an uncontrolled intersection. *Biometrika*, 49(1/2): 163–170.

Tassiulas, L.; and Ephremides, A. 1990. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. In *29th IEEE Conference on Decision and Control*, 2130–2132. IEEE.

Texas A&M Transportation Institute (TTI). 2021. Texas Transportation Institute. https://tti.tamu.edu/. Accessed: 2024-10-21.

Transport for New South Wales. 2024. SCATS Traffic System: Sydney Coordinated Adaptive Traffic System. https://scats.transport.nsw.gov.au/. Accessed: 2024-10-21.

Transport Research Laboratory (TRL). 2024. SCOOT: Split Cycle Offset Optimisation Technique. https://trl.co.uk/solutions/scoot. Accessed: 2024-10-21.

Troutbeck, R.; and Brilon, W. 1996. Unsignalized Intersection Theory. in Traffic Flow Theory. *Washington, DC: US Federal Highway Administration*, 8–1.

Underwood, R. 1961. Speed, Volume, and Density Relationship. Quality and Theory of Traffic Flow, Yale Bur.

Vrbanić, F.; Ivanjko, E.; Mandžuka, S.; and Miletić, M. 2021. Reinforcement learning based variable speed limit control for mixed traffic flows. In *2021 29th Mediterranean Conference on Control and Automation (MED)*, 560–565. IEEE.

Wei, H.; Xu, N.; Zhang, H.; Zheng, G.; Zang, X.; Chen, C.; Zhang, W.; Zhu, Y.; Xu, K.; and Li, Z. 2019. Colight: Learning network-level cooperation for traffic signal control. In *Proceedings of the 28th ACM international conference on information and knowledge management*, 1913–1922.

Weidmann, U. 1993. Transport technology for pedestrians: transport-technical characteristics of pedestrian traffic, literature review. *IVT series*, 90.

Whitham, G. B. 2011. *Linear and nonlinear waves*. John Wiley & Sons.

Wiering, M. A.; et al. 2000. Multi-agent reinforcement learning for traffic light control. In *Machine Learning: Proceedings of the Seventeenth International Conference (ICML'2000)*, 1151–1158.

Yang, J.; Purevjav, A.-O.; and Li, S. 2020. The marginal cost of traffic congestion and road pricing: evidence from a natural experiment in Beijing. *American Economic Journal: Economic Policy*, 12(1): 418–453.

Yoo, J.-Y.; Sengul, C.; Merz, R.; and Kim, J. 2011. Experimental analysis of backpressure scheduling in IEEE 802.11 wireless mesh networks. In *2011 IEEE International Conference on Communications (ICC)*, 1–5. IEEE.

Zhu, F.; and Ukkusuri, S. V. 2014. Accounting for dynamic speed limit control in a stochastic traffic environment: A reinforcement learning approach. *Transportation research part C: emerging technologies*, 41: 30–47.