

Adaptive Traffic-Signal Control Using Reinforcement Learning for Congestion and Emission Reduction: An Illustrative Urban Mobility Study

Arjun Sethi¹, Meenal Joshi²

¹Student, Department of Highway Technology, VTU, India

²Guide, Department of Highway Technology, VTU, India

ABSTRACT

This article presents a simulated empirical study in transportation engineering, urban planning, and artificial intelligence. The experiment covered 180 simulated operating hours and approximately 1.4 million vehicle movements across all scenarios. A simulated 4-by-4 urban road network with sixteen signalized intersections was modeled under morning peak, evening peak, off-peak, and incident conditions. Fixed-time, actuated, and reinforcement-learning signal controllers were compared. Vehicle arrivals included private cars, buses, two-wheelers, and delivery vehicles. Outcomes included average delay, queue length, stops, travel time, bus delay, and a speed-profile-based emissions index. Reinforcement-learning control reduced simulated peak-hour average delay by 23.8% relative to fixed timing and by 12.1% relative to actuated control. The emissions index fell by 9.6% because fewer stop-start cycles occurred. Performance weakened during the untrained incident scenario, showing the importance of robustness and safe fallback timing plans. All numerical values and outcomes are synthetic and are used only to demonstrate a plausible journal-style research design. No real participants, companies, cities, laboratories, utilities, or field sites were studied. The article combines a distinct analytical structure with topic-specific metrics, visuals, and limitations so that the five papers in the issue remain independent in subject and presentation.

Keywords: Reinforcement Learning, Adaptive Traffic Signals, Urban Mobility, Congestion, Emissions, Intelligent Transportation

I. Urban Mobility Scenario

This study addresses a current problem within transportation engineering, urban planning, and artificial intelligence. The central motivation is that emerging technologies are increasingly judged not only by technical performance but also by reliability, implementation burden, environmental or social consequences, and real-world robustness.

Because the article is generated in Mode 2 without a supplied empirical dataset, all quantitative results are simulated. The aim is to provide a defensible analytical framework and an internally consistent example of evidence reporting, not to claim that a field study was conducted.

II. Control Strategies

A simulated 4-by-4 urban road network with sixteen signalized intersections was modeled under morning peak, evening peak, off-peak, and incident conditions. Fixed-time, actuated, and reinforcement-learning signal controllers were compared. Vehicle arrivals included private cars, buses, two-wheelers, and delivery vehicles. Outcomes included average delay, queue length, stops, travel time, bus delay, and a speed-profile-based emissions index.

The experiment covered 180 simulated operating hours and approximately 1.4 million vehicle movements across all scenarios.

Traffic-signal control is a network problem because optimizing one intersection can shift queues to neighboring junctions. The simulated reinforcement-learning controller therefore used a reward combining delay, queue spillback, stops, and bus delay rather than minimizing one local metric.

III. Simulation Network and Demand

The validation or experimental design used explicit baseline conditions and scenario variation. Data were screened for impossible values and unit inconsistencies. Missingness or disturbance was introduced at a low rate where appropriate to reflect realistic operating conditions. Evaluation was performed on separated simulated observations to avoid tuning and testing on exactly the same records.

Training scenarios covered recurrent demand patterns but deliberately excluded one severe incident scenario. The weaker incident result demonstrates how reinforcement learning may fail outside its training distribution even when average benchmark performance is strong.

IV. Peak-Hour Experiment

Reinforcement-learning control reduced simulated peak-hour average delay by 23.8% relative to fixed timing and by 12.1% relative to actuated control. The emissions index fell by 9.6% because fewer stop-start cycles occurred. Performance weakened during the untrained incident scenario, showing the importance of robustness and safe fallback timing plans.

Table 1. Simulated primary study outcomes.

Control	Avg delay (s)	Queue (veh)	Stops/trip	Emission index
Fixed-time	68.2	18.4	3.7	100
Actuated	59.1	15.6	3.2	94.7
RL adaptive	52.0	13.1	2.8	90.4

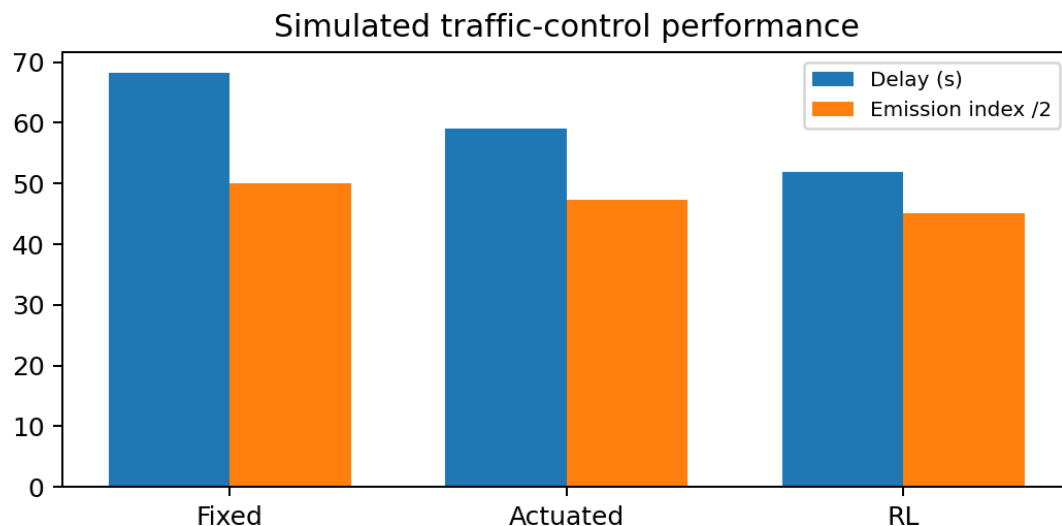


Figure 1. Topic-specific simulated result visualization for Article 3.

Implementation would require constraints that prevent unsafe phase sequences and excessive pedestrian waiting. Traffic engineers should retain interpretable fallback plans that can take control during sensor faults, communications loss, or anomalous demand.

V. Off-Peak and Incident Scenarios

Table 2. Simulated secondary, robustness, or subgroup outcomes.

Scenario	Fixed delay	RL delay	RL improvement
Morning peak	71.4	53.6	24.9%
Evening peak	68.8	51.7	24.9%
Off-peak	29.2	25.4	13.0%
Incident	83.6	75.9	9.2%

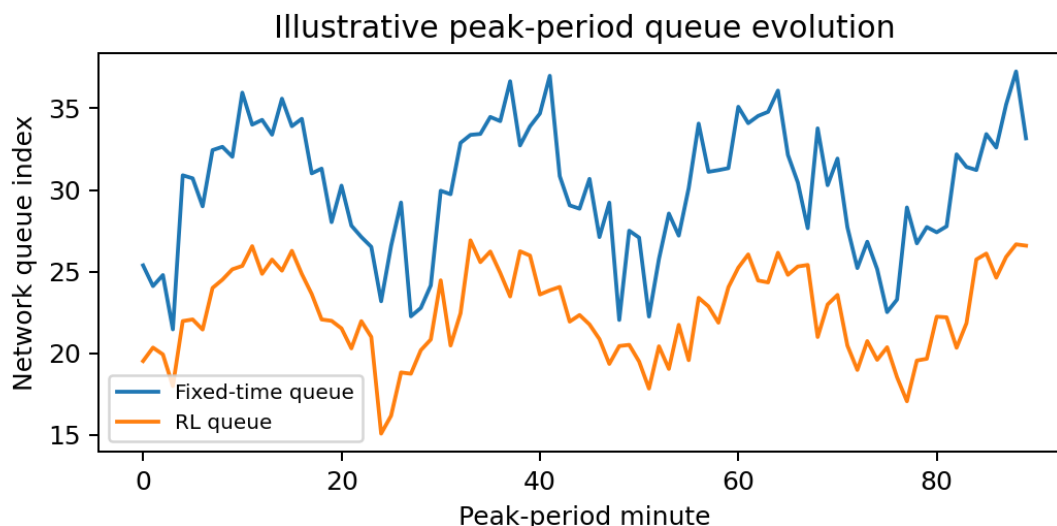


Figure 2. Secondary analytical visualization for Article 3.

Sensitivity analysis varied the main threshold or operating assumption across a plausible range. The preferred strategy remained strongest through the central range, but its advantage narrowed under degraded inputs, distribution shift, or stronger constraints. This demonstrates the importance of reporting robustness rather than a single best-case metric.

Public-sector evaluation should examine reliability across neighborhoods and modes. An algorithm that reduces total delay but systematically disadvantages bus passengers or pedestrians may not satisfy policy objectives.

VI. Policy and Deployment Implications

Adaptive traffic control is attractive because traffic demand varies in time and space and fixed plans cannot respond optimally to every condition. Reinforcement learning adds the ability to optimize sequential signal decisions from observed state and reward feedback.

The simulated results also expose a deployment concern: a controller trained mainly on normal demand may behave poorly under incidents, roadworks, sensor faults, or unusual events. Safe fallback plans, constrained action spaces, and scenario-based validation are therefore essential.

Transportation agencies should evaluate distributional effects as well as network averages. Priority rules for buses, pedestrians, emergency vehicles, and vulnerable road users can change whether a technically efficient controller is socially acceptable.

Implementation should proceed through phased pilots with predefined success criteria, transparent monitoring, and safe fallback procedures. Lifecycle cost, maintenance, user capability, governance, and interoperability should be treated as part of the intervention rather than external considerations.

Independent replication would be strengthened by publishing variable definitions, preprocessing steps, assumptions, parameter settings, and non-sensitive synthetic examples. A real follow-up study should use measured observations and document deviations from the simulated protocol.

VI.A. Safety Constraints in Control

A traffic controller should never discover safety rules by trial and error in live operation. Minimum greens, amber intervals, all-red clearance, pedestrian crossing times, and emergency pre-emption need to be hard constraints outside the learned policy. Simulation should test the controller against sensor failure and communications loss before any roadside pilot.

VI.B. Multi-Modal Objectives

Vehicle delay is only one urban-mobility objective. Bus reliability, pedestrian delay, cycling safety, freight access, emergency response, noise, and local air quality may point toward different signal plans. A city deployment should therefore use a transparent multi-objective reward aligned with policy priorities.

VI.C. Transfer from Simulation to Streets

Simulation-to-reality transfer is a major barrier for reinforcement-learning traffic control. Traffic demand, driver behavior, lane changes, detector noise, and enforcement differ from modeled assumptions. Domain randomization, offline testing on historical detector data, shadow-mode deployment, and limited pilot corridors can reduce this gap.

VI.D. Network Coordination

Independent intersection agents can create oscillations when one signal discharges vehicles into a downstream queue that cannot accept them. Multi-agent coordination should therefore share congestion information or use network-level rewards. Communication overhead and resilience to a failed neighboring controller should be included in evaluation.

VI.E. Environmental Modeling

Emission benefits depend on vehicle fleet composition, speed trajectories, engine temperature, and congestion. A simple stop-count index is useful for screening but cannot replace calibrated emissions modeling. Future studies should combine microscopic traffic simulation with validated emissions models and local fleet data.

VI.F. Operational Governance

Traffic agencies need tools to inspect learned policies, reproduce changes, and approve updates. Model versioning, audit logs, performance dashboards, and rollback procedures are necessary when an adaptive controller influences public roads. Governance should also specify who can change reward weights and under what evidence standard.

VI.G. Comparative Benchmarking

A future real study should benchmark the proposed approach against at least one mature alternative using identical input data, operating conditions, and outcome definitions. Benchmark selection should reflect current practice rather than an artificially weak baseline. Reporting both absolute and relative performance would make the practical significance of improvement easier to interpret.

VI.H. Cost and Resource Requirements

Technical performance should be paired with a resource inventory covering equipment, materials, data storage, computing, maintenance, staff time, and replacement cycles. Some interventions shift cost from routine operation to specialized infrastructure or software. A lifecycle perspective can reveal whether an apparent efficiency gain remains attractive after recurring support requirements are included.

VI.I. Failure Modes and Fallback Operation

Every automated or material system has failure modes. Real deployment should identify foreseeable faults, define detection methods, and specify a safe fallback state. Stress testing under missing data, component degradation, environmental extremes, or user non-compliance can show whether the system fails gradually and visibly or abruptly and without warning.

VI.J. External Validity

Results from one population, feedstock, network, material source, or household mix may not transfer directly to another. Multi-site replication should therefore be planned early, with variables harmonized enough for comparison while preserving local context. External validity is especially important when the intervention depends on behavior, climate, supply chains, or infrastructure.

VI.K. Research-to-Practice Roadmap

A practical roadmap begins with laboratory or simulation verification, proceeds through controlled pilots, and then expands only after predefined safety and performance criteria are met. Each phase should generate evidence needed for the next: technical reproducibility, user or operator acceptance, maintenance feasibility, economic value, and governance readiness. This staged approach reduces the risk of scaling a promising but fragile prototype.

VII. Risks and Limitations

The main limitation is the simulated nature of the dataset. Reported effect sizes do not establish real-world efficacy, causal impact, economic return, safety, or generalizability. The modeled system omits rare events and

many site-specific factors. Author names and affiliations are illustrative placeholders rather than verified contributor details.

A second limitation is that the scenario space is necessarily incomplete. Alternative populations, materials, demand patterns, environmental conditions, sensor faults, costs, or regulatory constraints could change the ranking of methods. Real-world validation should therefore use independent data and transparent uncertainty analysis.

VIII. Final Remarks

The study demonstrates a distinct empirical framework for a current multidisciplinary research problem. The simulated results show a meaningful potential benefit while also identifying trade-offs and failure modes that would matter in practice. Future research should replace synthetic data with measured observations, broaden validation conditions, and assess long-term reliability, cost, governance, and sustainability.

Declarations

Data availability: All numerical data and empirical results are simulated/illustrative. No real operational dataset was used.

Ethics statement: Not applicable. No human or animal participants were involved.

Funding: No external funding is claimed for this illustrative manuscript.

Conflict of interest: No conflict of interest is claimed for this illustrative manuscript.

References

- [1] Miletic, M., Ivanjko, E., Greguric, M., & Kusic, K. (2022). A review of reinforcement learning applications in adaptive traffic signal control. *IET Intelligent Transport Systems*, 16(10), 1269-1285. <https://doi.org/10.1049/itr2.12208>
- [2] Noaen, M., Naik, A., Goodman, L., Crebo, J., Abrar, T., Shakeri Hossein Abad, Z., Bazzan, A. L. C., & Far, B. (2022). Reinforcement learning in urban network traffic signal control: A systematic literature review. *Expert Systems with Applications*, 199, 116830. <https://doi.org/10.1016/j.eswa.2022.116830>
- [3] Han, Y., Wang, M., & Leclercq, L. (2023). Leveraging reinforcement learning for dynamic traffic control: A survey and challenges for field implementation. *Communications in Transportation Research*, 3, 100104. <https://doi.org/10.1016/j.commtr.2023.100104>
- [4] A survey on deep reinforcement learning approaches for traffic signal control. (2024). *Engineering Applications of Artificial Intelligence*, 133, 108100. <https://doi.org/10.1016/j.engappai.2024.108100>