future transportation

MDPI

Article

Hybrid Model-Based Framework for Real-Time Adaptive Traffic Signal Control

Bratislav Lukić [1], Goran Petrović [1], Žarko Ćojbašić [1], Dragan Marinković [2,3,4,*] and Srđan Dimić [5]

[1] Faculty of Mechanical Engineering, University of Niš, 18000 Niš, Serbia; bratislav.lukic@masfak.ni.ac.rs (B.L.); goran.petrovic@masfak.ni.ac.rs (G.P.); zcojba@ni.ac.rs (Ž.Ć.)
[2] Faculty of Transport Systems and Mechanical Engineering, Technical University Berlin, 10623 Berlin, Germany
[3] University College, Korea University, Seoul 02841, Republic of Korea
[4] Institute of Mechanical Science, Vilnius Gediminas Technical University, 10105 Vilnius, Lithuania
[5] Military Academy, University of Defence in Belgrade, 11000 Belgrade, Serbia; srdjan.dimic@mod.gov.rs
* Correspondence: dragan.marinkovic@tu-berlin.de

Abstract

Real-time traffic signal control represents a key challenge in modern intelligent transportation systems, particularly under highly variable traffic flows and the presence of priority vehicles. This study proposes a hybrid framework for adaptive signal plan control at a signalized intersection. The framework integrates deep learning-based traffic prediction, surrogate-based performance evaluation, and reinforcement learning-based adaptive control. Short-term traffic flow is predicted using recurrent neural networks, providing anticipatory information for traffic control decisions. Based on predicted flows and generated candidate signal plans, a machine learning surrogate model enables fast estimation of key performance indicators, including average vehicle delay and queue length. Adaptive control is implemented using the Proximal Policy Optimization algorithm within the SUMO environment via TraCI, which enables real-time fine-tuning of signal phases. A dedicated priority and stability module ensures effective emergency vehicle preemption and adaptive public transport priority while preserving intersection stability. Simulation results show that the proposed framework reduces average vehicle delay by up to 35% compared with FT and by up to 15% compared with standalone RL, while also improving traffic flow efficiency and priority vehicle performance.

Keywords: deep learning; reinforcement learning; surrogate modeling; priority vehicle preemption; SUMO

Check for updates

Academic Editors: Wei Zhou, Weijie Yu and Liyang Hu

Received: 30 March 2026
Revised: 28 April 2026
Accepted: 29 April 2026
Published: 1 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Urban traffic systems face increasing challenges due to traffic growth, variability in flows, and stricter requirements related to safety, energy efficiency, and environmental impact. Traffic signal control (TSC) at intersections plays a key role in traffic management, as it directly affects traffic efficiency, average vehicle delay, queue length, and overall urban mobility.

Modern urban traffic systems are nonlinear and stochastic, with frequent changes in operating conditions that exceed the assumptions of traditional control strategies. These challenges are particularly evident under high-demand and oversaturated conditions, as well as in scenarios involving priority vehicles and multimodal interactions.

One of the key challenges in TSC is balancing efficiency and stability. Highly adaptive strategies may reduce average vehicle delays but often increase control variability and

instability. Conversely, more stable strategies may fail to respond adequately to dynamic traffic conditions.

Another critical challenge is anticipatory decision-making. Effective TSC depends on both the current traffic state and the ability to anticipate short-term changes and adjust signal plans accordingly. Systems without predictive capability react too late, which further degrades performance under heavy congestion. At the same time, incorporating prediction without proper optimization and adaptation mechanisms can result in suboptimal or unstable control decisions.

These challenges highlight the need for modern TSC systems to adopt an integrated approach that simultaneously combines prediction of future traffic states, globally optimized planning of signal parameters, and real-time adaptation. Such an approach must be robust under saturated conditions while ensuring reliable prioritization of emergency vehicles (EV) and public transport vehicles (PT) without destabilizing the system.

Motivated by the aforementioned challenges, this study presents a hybrid (HYB) framework, which integrates predictive, optimization, and adaptive components into a unified control framework. The goal of the proposed HYB framework is to overcome the limitations of individual control strategies through their complementary integration, thereby enabling anticipatory, stable, and efficient intersection control across a wide range of traffic conditions. Compared to single-paradigm control strategies, the system seeks to balance average vehicle delay reduction, control stability, and reliable priority implementation.

The main contribution of this paper is in the development and evaluation of the proposed HYB framework for TSC, tailored to the requirements of modern urban intersections. Through a detailed simulation-based analysis, the proposed HYB framework is shown to outperform conventional and standalone intelligent control strategies, particularly under high traffic demand and complex conflict scenarios. As a result, this work contributes to the development of robust and scalable solutions for urban TSC in real-world urban environments.

It should be noted that environmental performance indicators such as CO_2 emissions are not explicitly considered in this study due to the lack of reliable vehicle-level data required for accurate estimation. Therefore, the analysis focuses on operational performance indicators, including delay, queue length, priority performance, and control stability. The inclusion of emission-related metrics is left for future research.

2. Literature Review

The optimization of TSC at urban intersections is a key focus within Intelligent Transportation Systems (ITS), due to its direct effects on traffic efficiency, safety, emissions, and overall urban quality of life. Traditional static and semi-adaptive control strategies rely on predefined cycles, signal plans, and periodic adjustments based on detector data. However, these approaches often underperform in the face of high temporal and spatial variability in traffic flows and specific events such as EV and PT arrivals [1]. Their main drawback is their reactive nature, which limits their ability to predict future system states.

Classical adaptive systems, such as SCOOT, SCATS, and RHODES, marked a significant technological leap by allowing automated adjustments to phase durations and cycle lengths based on current traffic conditions [2–4]. Despite their long-term use, subsequent studies show these systems struggle with the complex, multimodal, and highly nonlinear traffic patterns typical of modern cities. This is especially evident during unpredictable congestion, vehicle–pedestrian interactions, and dynamic priority management. Another widely used practical approach is the actuated TSC, which dynamically extends or terminates signal phases based on detector inputs and gap-out conditions, enabling the controller to respond to real-time traffic demand at individual approaches [5].

The max-pressure (MP) control strategy, initially proposed by [6], is a decentralized adaptive TSC. It selects the phase with the highest "pressure", which is defined as the difference between upstream and downstream queues. It is well-regarded for its strong theoretical guarantees concerning network stability and throughput maximization. Due to these advantages and its relatively straightforward implementation, the MP is often used as a benchmark in studies of adaptive TSC and RL-based control strategies.

With the growing availability of large volumes of traffic data, the past decade has seen substantial advancements in machine learning (ML) and deep learning (DL) control strategies for traffic flow modeling and prediction. Deep neural networks, particularly long short-term memory (LSTM) architectures, effectively capture temporal dependencies and provide more accurate predictions than traditional statistical and shallow ML models [7]. This capability is further supported by the successful application of ML and DL techniques in TSC and ITS [8,9]. In addition, similar methodologies have demonstrated strong performance in other complex engineering domains characterized by nonlinear relationships, high-dimensional data structures, and multiscale interactions [10]. For instance, the study in [10] applies ML to model complex structure–property relationships in plastic deformation processes, illustrating the ability of ML to capture nonlinear dependencies in systems with high variability. Such characteristics are also present in traffic flow dynamics, where interactions between vehicles, TSC, and demand patterns exhibit similar complexity. Further improvements have been achieved through the development of spatio-temporal DL that jointly capture spatial and temporal dependencies in urban traffic networks [11–13]. Recently, research attention has increasingly shifted toward graph neural networks (GNNs) and transformer-based architectures, which enable more accurate modeling of complex interactions among intersections in traffic networks [14].

Alongside the development of predictive models, considerable progress has been made in adaptive TSC through the use of reinforcement learning (RL). Pioneering studies have shown that RL agents are capable of learning effective signal phase control policies by interacting with simulation environments [15,16]. Subsequent research introduced deep reinforcement learning (DRL) frameworks, leading to the development of architectures such as deep Q-network (DQN), dueling-DQN, proximal policy optimization (PPO), and advantage actor-critic (A2C), which have demonstrated substantial improvements in average vehicle delay reduction and traffic throughput [8,17].

Concurrently, the integration of predictive models with optimization techniques has been successfully explored in engineering applications, particularly through the combination of artificial neural networks and intelligent optimization methods [18]. For instance, the study in [19] combines artificial neural networks with fuzzy logic to model complex relationships and optimize process parameters in systems characterized by nonlinear interactions and uncertainty. These hybrid modeling–optimization approaches, originally developed in other engineering domains, are conceptually transferable to TSC as methodological analogues, since they address nonlinear system behavior, uncertainty, and complex decision spaces. Since 2018, DRL approaches have been further enhanced through the adoption of multi-agent systems, centralized–decentralized policies, and integration with predictive models [20–23]. Nevertheless, numerous studies show that RL and DRL systems often suffer from stability issues, high sensitivity to traffic fluctuations, increased performance variability, and excessive phase switching, especially under oversaturated conditions and in complex traffic flow conflicts.

In addition to prediction and adaptive control, an important line of research focuses on optimizing signal plans using metaheuristic optimization. Methods such as particle swarm optimization (PSO) [24] and grey wolf optimizer (GWO) [25] have demonstrated strong performance in addressing nonlinear and multimodal optimization problems, including the

determination of optimal phase durations and green time allocations. More recently, hybrid metaheuristic optimization that combines complementary search strategies have gained increasing attention, offering improved convergence properties and solution quality across complex optimization problems [26,27]. Despite these advantages, such approaches are typically applied in offline or quasi-online settings, which limits their ability to support continuous real-time adaptation. This limitation highlights the need for integrated frameworks that combine global optimization capabilities with real-time adaptive control. In this context, the study in [28] applies an improved whale optimization algorithm combined with an extreme learning machine to address fault diagnosis in mechanical systems characterized by nonlinear dynamics and high-dimensional feature spaces. These hybrid optimization frameworks, although originating from mechanical engineering applications, are considered methodological analogues relevant to TSC due to their ability to handle nonlinear dynamics, stochastic demand patterns, and high-dimensional optimization problems.

A comprehensive review indicates that single-paradigm approaches are insufficient to address the complexity of modern traffic systems. This applies to DL-based prediction, metaheuristic optimization, and RL/DRL-based control. While the DL-based prediction modules provide accurate predictions of future traffic states, they do not guarantee optimal control decisions. Metaheuristic methods yield globally optimized signal plans but lack adaptivity. Conversely, RL/DRL approaches enable real-time adaptation but often suffer from instability and limited anticipatory capability.

For these reasons, there is a clear need for a unified framework that simultaneously integrates predictive modeling, global optimization, and adaptive control, rather than treating these components as independent solutions. Such a proposed HYB framework enables proactive decision-making based on predicted traffic flows, globally optimized signal plans, and stable real-time adaptation. This is particularly important for the reliable implementation of priority strategies for EV and PT. In this context, the proposed HYB framework represents a direct synthesis of contemporary research directions by combining the DL-based prediction module for traffic flow prediction, a metaheuristic optimization module for generating candidate signal plans, and an RL-based control module for real-time fine-tuning and adaptive control. In this way, the model overcomes the limitations of individual approaches identified in the literature and constitutes a robust and scalable solution for TSC in complex urban environments.

3. Methodology and Model

This chapter presents the methodology of the proposed HYB framework for optimizing and adapting traffic signal plans at a signalized intersection. The proposed HYB framework combines the DL-based prediction module, metaheuristic optimization, supervised ML, RL, and a phased multi-criteria decision-making framework. It explicitly accounts for priority handling for EVs and PT, as well as the operational stability constraints of the intersection.

3.1. Architecture of the Proposed HYB Framework

The proposed HYB framework illustrated in Figure 1 is organized as a sequence of interconnected functional modules forming a unified control pipeline. The overall workflow starts from data acquisition and ends with real-time signal execution and stability supervision. The proposed HYB framework consists of the following modules:

- Input data acquisition and integration module, which collects and synchronizes heterogeneous traffic data from detectors, cameras, and simulation interfaces, providing a consistent state representation of the intersection;

- EV and PT detection and classification module, which detects priority vehicles and generates corresponding priority requests based on their type, location, and estimated time of arrival (ETA);
- DL-based traffic flow prediction module, which forecasts short-term traffic states (e.g., flows, queue lengths, and arrival rates) using historical and real-time data, enabling anticipatory control decisions;
- Metaheuristic optimization module, which generates a set of candidate signal plans based on predicted traffic conditions, optimizing key signal parameters such as cycle length, green splits, and phase sequence;
- ML-based surrogate evaluation module, which provides fast performance estimation of candidate signal plans, reducing the computational burden of detailed simulation-based evaluation;
- RL-based adaptive control module, which operates within the simulation of urban mobility (SUMO) environment and performs real-time fine-tuning of the selected signal plan through continuous interaction with the traffic system;
- Priority and stability management module, which enforces priority constraints for EV and PT while ensuring operational stability, preventing excessive phase switching and oscillatory control behavior.

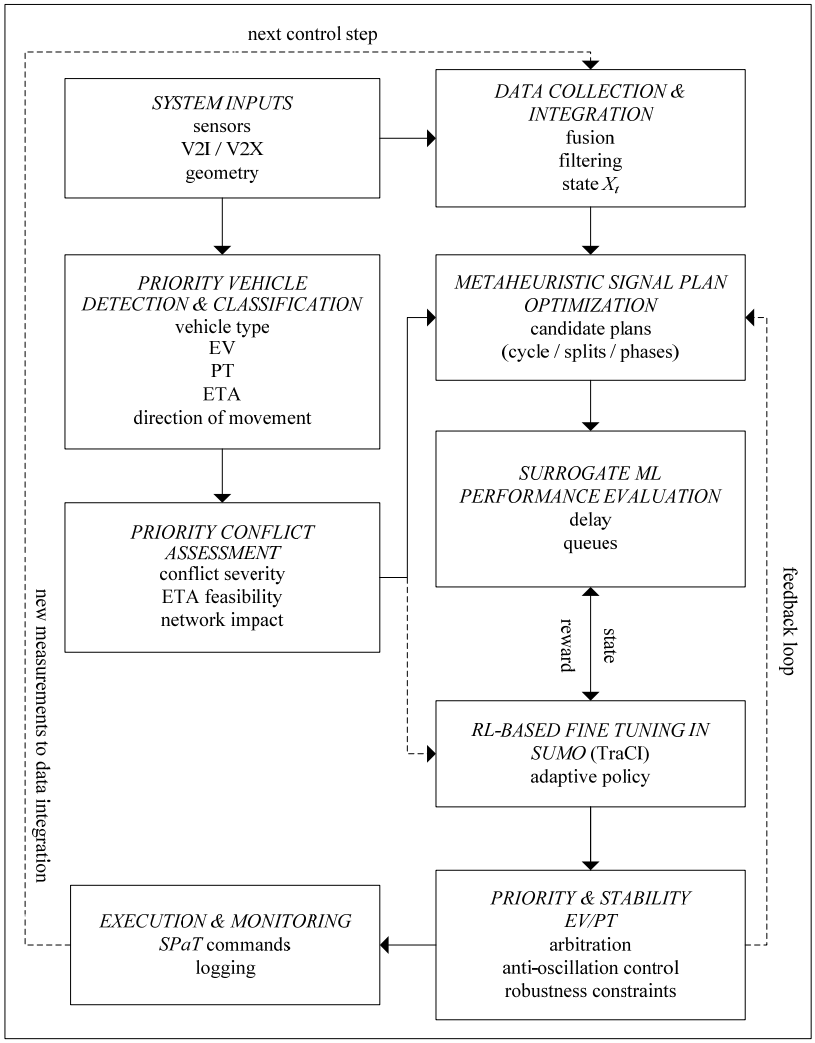


Figure 1. Block diagram of the proposed HYB framework.

The proposed HYB framework is a multi-layer system integrating prediction, optimization, and adaptive control. Each module has a defined role, and their interactions enable adaptation to dynamic traffic conditions.

To further clarify the computational workflow, the proposed HYB framework distinguishes between offline and online computational stages. Offline processes include model training, surrogate model construction, and metaheuristic signal plan generation. These tasks are executed prior to deployment or periodically due to their computational complexity. During online operation, lightweight computations are performed within each control cycle. Real-time traffic data are processed through the trained prediction model, and the RL agent selects control actions within a constrained action space. This separation ensures efficient real-time operation while maintaining high solution quality.

3.2. Input Data and Preprocessing

The input dataset used in this study comprises key traffic and operational parameters required for modeling and optimizing TSC. The observed data include traffic flow rates at individual approaches and lanes, as well as queue lengths and vehicle waiting times derived from simulation results. In addition, historical signal plans were used, including phase structure, green time durations, total cycle length, and intergreen times.

Prior to applying the proposed HYB framework, data preparation was performed using standard preprocessing procedures. These procedures include the identification and filtering out extreme values resulting from measurement errors or exceptional traffic conditions, interpolation of missing observations to maintain the continuity of time series, and normalization of input variables to ensure numerical stability and accelerate the learning process. In this way, a consistent and reliable input dataset was formed, suitable for the application of DL algorithms, metaheuristic optimization, and RL-based adaptive control.

3.3. EV/PT Detection and Classification

The detection and classification of priority vehicles represent a key component of adaptive TSC in modern ITS. Priority vehicles require timely identification to ensure safe and efficient passage through signalized intersections.

Contemporary approaches increasingly combine computer vision methods with communication technologies based on vehicle-to-infrastructure (V2I) and vehicle-to-everything (V2X) concepts [29]. Visual detection, implemented using deep convolutional neural networks, enables real-time identification of vehicle types, while V2I/V2X communication provides reliable information on vehicle position, speed, direction of travel, and priority level, even under conditions of limited visibility. By integrating these two approaches, higher robustness and detection accuracy are achieved.

In real-world implementations, the detection of EV can be achieved using advanced computer vision techniques. In particular, DL-based object detection models, such as the You Only Look Once (YOLO) family of algorithms, enable fast and accurate identification of specific vehicle classes from video streams in real-time [30,31]. These models are capable of recognizing EV based on visual features, including vehicle shape, color patterns, and the presence of distinctive markings or light signals. Such approaches have been successfully applied in ITS for real-time traffic monitoring and priority vehicle detection, demonstrating high detection accuracy and low latency. This makes them well-suited for integration into adaptive TSC frameworks. For PT, detection and data acquisition are typically performed using V2I or V2X communication technologies, which enable direct transmission of vehicle state information, including position, speed, and route, to the TSC infrastructure.

After the detection phase, vehicles are classified according to predefined priority levels, enabling differentiated control of signal phases. This information is forwarded to

the RL-based adaptive control module, where it is incorporated either into the system state or as an additional reward signal. Consequently, the RL agent learns control policies that explicitly favor higher-priority vehicles.

In this study, the detection and classification of priority vehicles are not based on physical sensors, but rather on simulation assumptions implemented within the SUMO environment. The detection and classification of priority vehicles are implemented using a combination of a simulation-based V2I approach and rule-based classification at the infrastructure level, in order to ensure reliable and reproducible evaluation of the proposed control algorithm. Within the SUMO simulation environment, priority vehicles are predefined using dedicated vehicle types and routes, while their presence, position, and motion are detected in real-time via the traffic control interface (TraCI) application programming interface (API). Vehicle identity, speed, distance to the intersection, and ETA are obtained via TraCI and included in the DRL agent's state representation.

Vehicle classification is performed based on predefined categories and a priority hierarchy, where EV are treated as having absolute priority, while PT are considered as relative priority. This approach enables a clear separation of the effects of priority requests from the baseline traffic flow, without introducing additional uncertainty associated with real-world visual detection systems.

3.4. Short-Term Traffic Flow Prediction Using DL

Within the proposed HYB framework for adaptive TSC, short-term traffic flow prediction represents a central input module that provides anticipatory information for the optimization and control components of the system. This module is directly integrated into the architecture shown in Figure 1, where prediction results are used to construct the system state and enhance decision-making in subsequent time steps.

Traffic flow prediction is implemented using recurrent neural networks, specifically LSTM and GRU architectures, which are well suited for modeling nonlinear time series with pronounced short-term and medium-term dependencies [32]. These models enable accurate learning of temporal dynamics in traffic flow intensities at individual intersection approaches.

Let $q_i(t)$ denote the vehicle flow rate at approach i at time instant t, while vector (1) represents the input feature vector composed of historical traffic flow observations:

$$q(t) = \{q_1(t), q_2(t), \ldots, q_m(t)\} \tag{1}$$

The objective of the prediction module is to estimate the future traffic flow intensity $\hat{q}_i(t+k)$ for a predefined prediction horizon k. The prediction model can be formally expressed as (2):

$$\hat{q}(t+k) = f_\theta(q(t-n+1), \ldots, q(t), x(t-n+1), \ldots, x(t)) \tag{2}$$

where f_θ denotes a nonlinear mapping implemented by an LSTM/GRU neural network with parameters θ, while n represents the length of the input temporal window, and $x(t)$ denotes additional input variables, including queue lengths, vehicle waiting times, information on active signal phases, and indicators of the presence of priority vehicles.

In this study, the input temporal window is set to $n = 12$ time steps, corresponding to a 60 s historical observation window, while the prediction horizon is set to $k = 3$ time steps, corresponding to a 15 s prediction horizon. The selected GRU architecture consists of two stacked recurrent layers with 64 hidden units each, followed by a fully connected output layer. Dropout regularization with a rate of 0.20 is applied between the recurrent layers to reduce overfitting. The model is trained using the Adam optimizer with a learning rate

of 0.001 and weight decay of 10^{-5}. The loss function is defined as the mean squared error (*MSE*) between predicted and actual traffic flows (3):

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (3)$$

where y_i and $\hat{y}_i$ denote the observed and predicted values, respectively, and N is the number of samples.

Training is performed using mini batches of size 64 over a maximum of 80 epochs. Early stopping is applied with a patience of 10 epochs to prevent overfitting, and gradient clipping is used to ensure numerical stability. Additionally, a ReduceLROnPlateau learning-rate scheduler is employed to adapt the learning rate during training. Model validation is conducted using a chronological data split to preserve temporal consistency. The dataset is divided into 70% training data, 15% validation data, and 15% test data, without random shuffling. Prediction performance is evaluated using multiple metrics, including mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). The prediction model was evaluated on the test dataset, and the obtained results confirm high forecasting accuracy across all considered metrics, as summarized in Table 1. These results indicate that the prediction module provides reliable input for the subsequent optimization and control components of the proposed HYB framework.

Table 1. Prediction model performance.

Metric	MAE	RMSE	MAPE	R^2
Value	17.6	24.9	9.3	0.93

The outputs of the prediction module are directly used by the metaheuristic optimization module for generating candidate signal plans and by the RL-based control module as part of the system state representation. In this way, the predictive component enables anticipatory control actions, reduces system reactivity, and contributes to improved stability and efficiency of the proposed HYB framework.

3.5. Metaheuristic Optimization of Signal Plans

Based on the predicted traffic flows, the metaheuristic optimization module formulates and evaluates a set of candidate signal plans and selects the optimal solution with respect to traffic efficiency, priority requirements of EV and PT, and control stability. The optimization problem is formulated as a constrained multi-variable problem, where the decision vector is defined as (4):

$$S = [C, g_1, g_2, \ldots, g_m, \phi_1, \phi_2, \ldots, \phi_m] \quad (4)$$

where C denotes the cycle length, g_j represents the green time of phase j, and ϕ_j denotes the phase offset or phase-order parameter.

To solve this problem, a hybrid metaheuristic algorithm combining PSO and GWO is employed. The PSO component provides efficient global exploration through velocity-based updates, while the GWO component enhances exploitation by simulating leadership hierarchy and encircling behavior.

The hybridization is implemented through a convex combination of PSO and GWO updates (5):

$$S^{new} = \lambda \cdot S^{PSO} + (1 - \lambda) \cdot S^{GWO} \quad (5)$$

The PSO parameters are defined as inertia weight $w = 0.7$, cognitive coefficient $c_1 = 1.5$, and social coefficient $c_2 = 1.5$. The GWO component uses a linearly decreasing coefficient $a \in [2,\ 0]$ to transition from exploration to exploitation over iterations.

The optimization is performed with a population size of 30 candidate solutions over 50 iterations. These values provide a balance between solution quality and computational efficiency, which is consistent with the real-time applicability requirements of the proposed HYB framework.

The optimization is subject to the following constraints (6):

$$\begin{gathered} \sum_{j=1}^{m} g_j + L \leq C \\ g_j^{min} \leq g_j \leq g_j^{max},\ \forall j \\ C^{min} \leq C \leq C^{max} \end{gathered} \tag{6}$$

where L represents the total lost time per cycle.

This study adopts a hybrid PSO-GWO algorithm, which combines the global search capability of the GWO with the local exploitation ability of PSO, making it particularly suitable for TSC problems involving combined discrete and continuous constraints, as well as stability requirements. Similar hybrid metaheuristic optimization has been widely adopted in recent studies to enhance the exploration–exploitation balance in complex optimization problems [33].

Unlike conventional hybrid metaheuristics that focus solely on standalone optimization, the proposed control strategy integrates the optimization module directly into a hierarchical control framework. Within this structure, the initial signal plans are further refined through RL. This interaction between metaheuristic optimization and RL represents a key contribution of the proposed HYB framework, as it enables structured reduction of the action space, improves convergence stability, and enhances control robustness.

To ensure feasibility, a repair mechanism is applied after each iteration. It clamps cycle length and green times within bounds, rescales green times if needed, enforces cycle stability, and adjusts phase offsets.

The population of signal plans is iteratively updated while respecting constraints on minimum and maximum green times, cycle length, and stability (maximum allowable deviation from the nominal cycle). The objective function is defined as (7):

$$F(S) = w_1 \cdot D_{tav}(S) + w_2 \cdot Q_{tav}(S) + w_3 \cdot D_{EV}(S) + w_4 \cdot D_{PT}(S) + w_5 \cdot (1 - Stab(S)) \tag{7}$$

where the weighting coefficients satisfy $w_3 = 0.30$, $w_4 = 0.20$, $w_1 = 0.20$, $w_2 = 0.15$, and $w_5 = 0.15$ and where

- D_{tav}—Total average vehicle delay;
- Q_{tav}—Total queue length;
- D_{EV}—EV delay;
- D_{PT}—PT delay;
- $Stab(S) \in \{0, 1\}$, (1 = fully stable).

The weighting coefficients were determined through empirical tuning based on preliminary simulation experiments and validated through sensitivity analysis to prioritize EV and PT performance while maintaining overall traffic efficiency and control stability.

From the perspective of the overall control architecture, the metaheuristic optimization module also plays an important role in structuring the decision space for the RL-based control module (8):

$$\pi^* = arg \min_{\pi \in \Pi} \mathbb{E}\left[\sum_{t=0}^{T} D(s_t, a_t)\right] \tag{8}$$

where s_t denotes the system state at time t, a_t represents the control action, and $D(s_t, a_t)$ denotes the cost function and $\mathbb{E}[\cdot]$ represents the expectation operator. In this formulation, the objective is to minimize traffic-related costs.

However, when the action space A is large, the learning process may become unstable and less efficient. In the proposed HYB framework, the metaheuristic optimization module generates a constrained set of feasible signal plans, effectively reducing the RL action space $A' \subseteq A$. As a result, the RL-based control module performs local policy refinement within a structured solution space, which improves convergence stability and reduces oscillatory signal switching behavior. Such improvements in convergence behavior and solution stability are consistent with recent developments in hybrid metaheuristic optimization [34].

3.6. ML Surrogate Model for Performance Assessment

Since the metaheuristic optimization module may generate a large number of candidate signal plans, evaluating each plan through full SUMO microsimulation (version 1.19.0) via the TraCI interface leads to substantial computational overhead, while the ML and RL components are implemented in Python 3.10 using TensorFlow and XGBoost libraries. Recent studies in TSC and ITS have shown that data-driven surrogate models can effectively approximate simulation-based evaluation while maintaining high decision quality and substantially reducing computational cost [35,36]. Similar approaches have also been successfully applied in other engineering domains [37]. These findings demonstrate that ML and DL can replace expensive numerical procedures in complex systems, which is consistent with the proposed surrogate-based evaluation of TSC strategies.

To address this challenge, the proposed HYB framework integrates an ML-based surrogate model that approximates simulation outputs and enables efficient evaluation of candidate solutions prior to adaptive control.

The surrogate model is trained using the results of previously executed SUMO microsimulations. For a given signal plan S and the predicted traffic flow vector $\hat{q}$, the surrogate model estimates the performance vector describing the expected traffic conditions at the intersection. This relationship can be expressed as (9):

$$y = g_\beta(S, \hat{q}) = \{D_{tav}, Q_{tav}, D_{EV}, D_{PT}\} \quad (9)$$

Within this study, several supervised learning techniques were evaluated, including Random Forest, XGBoost (version 1.7.6), and CatBoost (version 1.2.2) regression models, due to their ability to approximate nonlinear relationships and their robustness in handling heterogeneous traffic data.

The training dataset comprises 3600 input–output pairs generated from SUMO simulations across multiple traffic scenarios with varying demand levels, signal plans, and random initial conditions, ensuring sufficient diversity of traffic states and control configurations. The dataset is randomly split into training (70%), validation (15%), and test (15%) subsets. A comparative evaluation of Random Forest, CatBoost, and XGBoost models was performed using RMSE, MAE, R^2, and inference time, with the results summarized in Table 2.

Table 2. Comparative performance of surrogate models.

Model	RMSE	MAE	R^2	Inference Time (ms)
Random Forest	7.82	5.94	0.89	4.2
CatBoost	7.35	5.61	0.91	5.8
XGBoost	6.91	5.22	0.93	3.7

XGBoost demonstrated the best trade-off between prediction accuracy and computational efficiency and was therefore selected as the final surrogate model g_β. The corresponding hyperparameters are summarized in Table 3. The surrogate model is trained using input–output mappings (10).

$$(S, \hat{q}) \rightarrow y \tag{10}$$

Thereby enabling efficient ranking and filtering of candidate signal plans before they are passed to the adaptive control module. The RL-based control module subsequently operates on this filtered set of candidate signal plans, enabling real-time adaptive control within a constrained decision space.

Table 3. Implementation details and configuration of the proposed HYB framework.

Module	Component	Parameter	Value/Description
Prediction (DL)	Model type	Architecture	GRU (primary), LSTM (benchmark)
	Network structure	Layers	2 stacked recurrent layers
		Hidden units	64 per layer
		Dropout	0.20
	Input	Input window	12 time steps (60 s)
		Prediction horizon	3 time steps (15 s)
		Features	Flow, queue length, waiting time, signal phase, EV/PT indicators
	Training	Optimizer	Adam
		Learning rate	0.001
		Weight decay	1×10^{-5}
		Loss function	MSE
		Batch size	64
		Epochs	80 (maximum)
		Early stopping	Patience = 10
		Scheduler	ReduceLROnPlateau (factor 0.5, patience 5)
	Validation	Data split	70% train/15% validation/15% test
		Metrics	MAE, RMSE, MAPE, R^2
Metaheuristic Optimization	Algorithm type	–	Hybrid PSO-GWO
	Parameters	Population size	30
		Iterations	50
		PSO (w, c_1, c_2)	0.7, 1.5, 1.5
		GWO	α, β, δ hierarchy; parameter a decreases linearly ($2 \rightarrow 0$)
	Hybridization	Strategy	Convex combination of PSO and GWO updates
	Decision variables	–	Cycle length, green splits, phase sequence
	Constraints	–	$g_{min} \leq g \leq g_{max}$; $C_{min} \leq C \leq C_{max}$
	Stability	–	Maximum cycle deviation ΔC_{max}
	Objective function	–	Multi-criteria: average vehicle delay, queue length, EV delay, PT delay, stability
	Objective weights	w_1, w_2, w_3, w_4, w_5	0.20, 0.15, 0.30, 0.20, 0.15

Table 3. *Cont.*

Module	Component	Parameter	Value/Description
Surrogate Model (ML)	Models tested	–	Random Forest, XGBoost, CatBoost
	Selected model	–	XGBoost
	Input	–	Signal plan parameters + predicted traffic flows
	Output	–	Average vehicle delay, queue length, EV delay, PT delay
	Data	Dataset size	~3600 samples
		Split	70% train/15% validation/15% test
		Validation	Cross-validation (k = 5)
	Metrics	–	RMSE, MAE
	Selection criterion	–	Best trade-off between accuracy and computational efficiency
	XGBoost Configuration	Number of estimators	200
		Maximum tree depth	6
		Learning rate	0.1
		Subsample ratio	0.8
		Column sample (colsample_bytree)	0.8
		Regularization parameter (λ)	1.0
		Hyperparameter tuning	Grid search + cross-validation
RL (PPO)	Algorithm	–	Proximal Policy Optimization
	State space	–	Flow, queue lengths, signal phase, EV/PT indicators, predicted flows
	Action space	–	Top-ranked candidate signal plans
	Reward weights	w_1, w_2, w_3, w_4	0.25, 0.20, 0.30, 0.25
	Parameters	Learning rate	3×10^{-4}
		Discount factor (γ)	0.99
		Clip ratio (ε)	0.2
		Batch size	64
	Training	Episodes	500
		Steps per episode	300
		Optimization epochs	10
	Reward	–	Average vehicle delay, queue length, EV/PT priority
	Environment	–	SUMO + TraCI
Simulation Setup	Platform	–	SUMO
	Interface	–	TraCI
	Step length	–	1 s
	Runs per scenario	–	20 independent runs
	Warm-up period	–	~10 min (excluded from statistics)
	Traffic model	–	Krauss (car-following) + LC2013 (lane-changing)
Computational Setup	CPU	–	8-core/16-thread
	RAM	–	32 GB
	Storage	–	NVMe SSD
	GPU	–	Used for deep learning tasks

3.7. Adaptive Control Based on RL in the SUMO Environment

In the proposed HYB framework, signal plans ranked as the most efficient based on evaluations from the surrogate module are used as initial solutions for adaptive TSC using RL. The adaptive control module is implemented in a closed loop with the SUMO

simulation environment via TraCI, enabling bidirectional real-time information exchange between the control algorithm and the simulation.

At each discrete time step t, the system observes the state s(t), which includes current traffic flow rates, queue lengths, indicators of the presence of priority vehicles, the active signal phase, as well as short-term predicted flows obtained from the DL-based prediction module. The state representation is defined as a compact feature vector of dimension 20, ensuring a sufficient description of traffic conditions while maintaining computational efficiency.

Based on the observed state, an action $a(t)$ is selected by choosing one candidate signal plan from a constrained action space. Specifically, the action space is defined as a set of top-ranked candidate signal plans generated by the metaheuristic optimization module and filtered by the surrogate model. In this study, the action space is limited to 10 candidate signal plans, thereby significantly reducing the dimensionality of the control problem while preserving solution quality.

Adaptive control is implemented using the PPO algorithm, which is selected due to its learning stability, robustness to reward variability, and suitability for stochastic and real-time environments. The PPO agent is trained through continuous interaction with the SUMO environment, using policy updates based on clipped probability ratios to ensure stable learning.

The key hyperparameters of the PPO algorithm are defined as follows: learning rate $\alpha = 3 \times 10^{-4}$, discount factor $\gamma = 0.99$, clipping parameter $\epsilon = 0.2$, batch size of 64, and 10 optimization epochs per update. The training process is conducted over 500 episodes, with policy updates performed every 2000 simulation steps, corresponding to multiple episodes per update cycle.

The RL reward function is defined as (11):

$$R(t) = -(w_1 \cdot D_{tav}(t) + w_2 \cdot Q_{tav}(t) + w_3 \cdot D_{EV}(t) + w_4 \cdot D_{PT}(t)) \tag{11}$$

where $w_3 = 0.30,\ w_4 = 0.25, w_1 = 0.25,$ and $w_2 = 0.20$, indicating that the highest priority is assigned to EV, followed by PT. The training objective of the PPO agent is to maximize the expected cumulative reward (12):

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(t)\right] \tag{12}$$

where γ is the discount factor and $\gamma \in (0, 1)$.

The integration of the DL-based prediction module, surrogate-based performance evaluation, and PPO-based adaptive control leads to a significantly more stable and efficient learning process compared to conventional RL control strategies such as DQN. This improvement stems from the structured reduction of the state and action spaces. By limiting exploration to high-quality candidate solutions, the framework reduces reward variance and accelerates convergence.

All baseline control strategies (ACT, MP, and RL) are implemented under identical simulation conditions and evaluated at each simulation step (1 s) to ensure a consistent and fair comparison. The ACT controller follows standard detector-based logic, applying gap-out and max-out rules based on vehicle presence detected on each incoming lane. The MP controller selects the phase with the highest pressure, defined as the difference between upstream and downstream queue lengths for each movement, where downstream queues are estimated based on the occupancy of outgoing links; in the scenario of equal pressure values, a fixed phase order is used as a tie-breaking rule. The standalone RL controller uses a state representation including flow rates, queue lengths, signal phase, elapsed green

time, and EV/PT indicators. The action space consists of discrete TSC actions, including phase extension, phase termination, and phase switching, subject to the same operational constraints as the other control strategies. The reward function is formulated as a weighted combination of average vehicle delay, queue length, and priority-related terms, aligned with the objective of minimizing congestion while accounting for priority vehicles, but without the prediction and optimization components of the proposed HYB framework. All baseline control strategies are calibrated using consistent parameter settings and constraints to ensure a fair and unbiased comparative evaluation.

A detailed overview of the implementation settings, including model configurations, hyperparameters, and training procedures across all modules, is presented in Table 3.

Furthermore, the proposed hierarchical interaction between metaheuristic optimization and RL enhances control robustness by preventing excessive phase switching and ensuring stable signal timing transitions under varying traffic conditions.

3.8. Priority and Stability Control for EV/PT Vehicles

Within the proposed HYB framework, the priority and stability module is designed to ensure efficient passage of priority vehicles while simultaneously preserving the operational stability of the signalized intersection. Priority detection and activation are based on the following indicators for each approach i at time t:

- $p_i^{EV} \in \{0,1\}$—the presence of EV on the approach i;
- $p_i^{PT} \in \{0,1\}$—the presence of PT;
- $\Delta T_i^{PT} \in \{0,1\}$—indicator of PT delay relative to the timetable.

In the event that an EV is detected, a preemption mechanism is applied, temporarily assuming control over the signal plan. Preemption is implemented using one of the following strategies, depending on the current system state:

1. early activation of the green phase for the corresponding approach;
2. extension of the current green interval if the active phase is already serving the given approach;
3. activation of a dedicated EV phase with a limited duration.

These interventions are coordinated with the PPO-based control module, with preemption treated as a priority but time-limited action within the allowable control space.

For PT, an adaptive priority strategy is introduced, realized through limited extension of the green phase duration and, when necessary, modification of the phase sequence. Unlike preemption, PT priority does not disrupt the basic structure of the signal plan, but instead acts as a corrective mechanism within the adaptive RL control framework.

To preserve system stability, all priority interventions are subject to the following constraints (13):

$$\begin{gathered} |C(t) - C_0| \le \Delta C_{max} \\ N_{EV}(T_{win}) \le N_{EV}^{max} \\ \sum_{k \in T_{win}} \Delta g_{PT}^{(k)} \le G_{PT}^{max} \end{gathered} \tag{13}$$

where $C(t)$ is the current signal cycle length, C_0 is the nominal cycle time, $N_{EV}(T_{win})$ is the number of EV in the observed time window T_{win} and $\Delta g_{PT}^{(k)}$ represents the green phase extension for PT at step k.

Additionally, to prevent abrupt changes in traffic conditions, system dynamics constraints are introduced (14):

$$\begin{gathered} |Q_i(t) - Q_i(t-1)| \le \Delta Q_{max} \\ |D_i(t) - D_i(t-1)| \le \Delta D_{max} \end{gathered} \tag{14}$$

where $Q_i(t)$ and $D_i(t)$ represent the queue length and average vehicle delay on the approach i.

The PPO agent maximizes the RL reward function that penalizes average vehicle delay and queue length, while explicitly incorporating stability constraints, thereby ensuring EV/PT priority without significant degradation of overall system stability.

3.9. Execution and Monitoring

This module represents the final layer of the proposed HYB framework and ensures deterministic execution of control decisions. At each discrete time step t, the composite action $a(t)$ is mapped to a signal phase and timing (SPaT) command (15):

$$u(t) = \{\phi(t+1), g_\phi(t+1) \cdot \tau(t+1)\} \quad (15)$$

where $u(t)$ is the control action vector applied at decision step t, $\phi(t+1)$ is the index of the selected signal phase to be activated at the next time step $t+1$, $g_\phi(t+1)$ is the green time assigned to the selected phase ϕ at time step $t+1$, and $\tau(t+1)$ is a scaling factor used to adjust the effective duration of the green phase. This command is executed via the TraCI interface in the SUMO environment. After execution, the realized state is verified and, in the event of conflicts or safety constraints, a rollback to the last stable configuration is triggered.

The module continuously monitors EV/PT priority events and stability indicators. For analysis and reproducibility, the system state, control action, SPaT command, reward, and performance vector are logged at each step. This closes the control loop and ensures the operational coherence of the entire HYB.

The proposed HYB framework ensures real-time control by decoupling offline and online computational processes. Computationally intensive tasks—such as metaheuristic signal plan generation and model training—are handled offline or at extended intervals. In contrast, real-time operations rely on lightweight inference and constrained decision-making.

Preliminary evaluation indicates that the online control cycle, including state update, short-term traffic prediction, surrogate evaluation, and RL-based action selection, can be executed within sub-second time scales on standard computing hardware. Specifically, the average computation time per control cycle is 160 ± 22 ms, remaining well below 0.25 s on the specified hardware platform. This quantitatively confirms that the proposed HYB framework satisfies real-time requirements for intersection-level traffic signal control.

The modular design also enables potential scalability toward corridor-level and network-wide TSC. However, such extensions require additional validation, particularly with respect to inter-intersection coordination, spillback effects, and increased dimensionality of the state and action spaces, and are therefore considered as part of future work.

3.10. Statistical Analysis

A statistical evaluation framework was applied to assess the significance of differences between control strategies. Nonparametric tests were selected due to the non-normal distribution of traffic performance indicators. Pairwise comparisons were conducted using the Mann–Whitney U test, with Bonferroni correction applied to control for multiple comparisons. Effect sizes were calculated using rank-biserial correlation to quantify the magnitude of observed differences. The provided methodological descriptions, model configurations, and experimental settings are sufficient to enable independent reproduction of the proposed HYB framework.

4. Simulation Setup and Results

This section presents the results of applying the proposed HYB framework to a real-world scenario of an urban, heavily loaded isolated four-leg signalized intersection located in Vranje, Serbia. The experiments were conducted in the SUMO microsimulation environment.

The geometric configuration of the analyzed intersection, as illustrated in Figure 2, consists of four approaches, each with two traffic lanes. The length of each incoming link is approximately 300 m, while the effective storage length ranges between 80 m and 120 m depending on the approach. Lane lengths are consistent with the link geometry, and the speed limit is set to 50 km/h for all approaches, reflecting typical urban traffic conditions. On all approaches, the left lane is designated for through and left-turn movements, while the right lane is designated for through and right-turn movements. Permitted turning movements include through, left-turn, and right-turn maneuvers on all approaches, without dedicated protected left-turn phases, reflecting typical real-world operating conditions of urban signalized intersections.

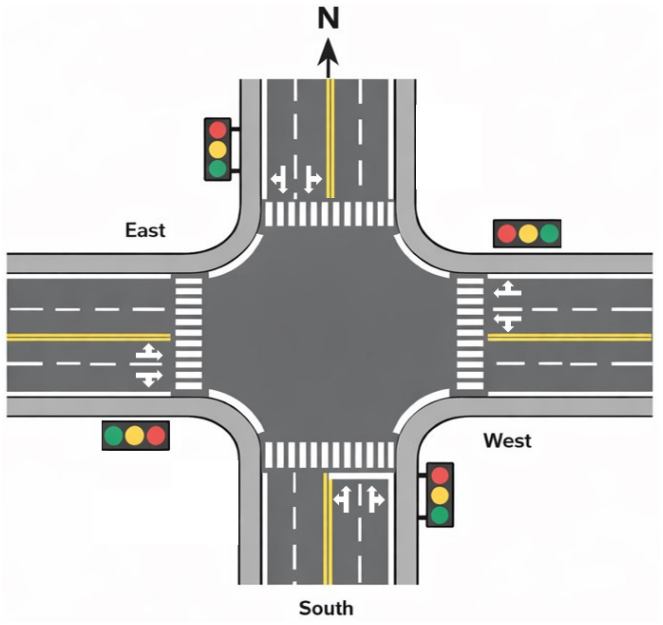


Figure 2. Schematic layout of the analyzed signalized intersection.

Turning movements follow the defined lane configuration and turning ratios. Vehicles share lanes depending on traffic demand. Interaction between conflicting movements is governed by TSC logic and right-of-way rules implemented in the SUMO simulation environment. As a result, turning vehicles may experience additional average vehicle delays due to opposing traffic flows and limited gap availability, particularly under higher demand levels. This modeling approach ensures a realistic representation of turning behavior and its impact on intersection performance.

TSC scheme is based on a two-phase plan, where opposing approaches are grouped within each phase. Intergreen times are defined in accordance with standard traffic engineering practice, including yellow and all-red intervals. Pedestrian crossings are provided on all approaches and are integrated into the corresponding signal phases.

Traffic demand levels were defined to represent four traffic demand groups (G1–G4), corresponding to increasing degrees of traffic saturation, including oversaturated conditions. The corresponding demand levels and turning ratios are summarized in Table 4. In addition, four operational scenarios (S1–S4) were defined. S1 represents the baseline scenario without priority, S2 corresponds to a priority scenario with the presence of EV,

and S3 represents a combined priority scenario including both EV and PT vehicles. Scenario S4 includes increased pedestrian demand and is introduced to reflect more complex multimodal traffic conditions. However, since the present study focuses on vehicle-based performance indicators, the evaluation and comparative analysis are limited to scenarios S1–S3.

Table 4. Traffic demand levels and turning ratios.

Group	Flow (veh/h)	Left (%)	Through (%)	Right (%)	Description
G1	300–400	20	60	20	Low demand (unsaturated)
G2	500–700	25	55	20	Moderate demand
G3	800–1000	30	50	20	High demand
G4	1100–1300	30	50	20	Oversaturated

Priority vehicle behavior is explicitly defined within the simulation environment through predefined vehicle types and corresponding control logic. EV are assigned absolute priority and trigger immediate TSC responses, including early phase activation and green extension when approaching the intersection. PT are assigned relative priority, where signal adjustments are applied in a controlled manner, such as limited green extension or phase reordering within predefined operational constraints.

In the simulation, priority vehicles adhere to standard microscopic driving behavior, governed by established car-following and lane-changing models. Their preferential treatment is implemented solely through TSC adaptation rather than alterations to vehicle dynamics. This approach ensures a consistent and realistic evaluation of priority strategies without introducing additional behavioral assumptions.

Vehicle arrivals were modeled using calibrated Poisson distributions, with arrival rates derived from field-measured traffic data, while detector behavior is represented through equivalent virtual sensing within the SUMO simulation environment. Traffic demand is categorized into four levels (G1–G4), representing low, moderate, high, and oversaturated conditions, implemented through calibrated arrival processes. Demand levels (G1–G4) are defined in Table 4, and all control strategies are evaluated across the full set of demand-scenario combinations.

This modeling approach is widely adopted in microscopic traffic simulation studies for urban environments. The calibration procedure involved matching simulated queue lengths and average vehicle delays with observed values by adjusting traffic demand parameters based on field data. Lane-changing behavior is governed by the LC2013 model implemented in SUMO. This approach was adopted to ensure consistency and fairness across all evaluated control strategies, as the primary objective of the study is to compare control logic rather than to reproduce site-specific driver behavior in detail.

Vehicle dynamics in the SUMO simulation environment are modeled using the Krauss car-following model [38] with explicitly defined parameter values. The main parameters are set as follows: desired time headway $\tau = 1.0$ s, driver imperfection $\sigma = 0.5$, maximum acceleration $a = 2.6$ m/s^2, comfortable deceleration $b = 4.5$ m/s^2, and minimum gap minGap = 2.5 m. Lane-changing behavior is governed by the LC2013 model, with parameters $lcStrategic = 1.0$, $lcCooperative = 1.0$, $lcSpeedGain = 1.0$ and $lcKeepRight = 0.5$. These values correspond to typical urban traffic conditions and are widely used in SUMO-based traffic simulation studies.

Lane-changing decisions in the LC2013 model are based on a combination of safety, incentive, and cooperation criteria. A lane change is performed only if sufficient space is

available in the target lane, ensuring collision-free behavior, while incentive conditions evaluate potential benefits such as reduced travel time or improved route alignment. Cooperative behavior is also considered, allowing surrounding vehicles to adjust their speed to facilitate lane changes. Under congested traffic conditions, these mechanisms limit excessive lane-changing activity, as reduced gaps and increased vehicle density constrain feasible maneuvers.

Under oversaturated traffic conditions, lane discipline is governed by the interaction of car-following and lane-changing models. These mechanisms jointly enforce safety constraints, gap acceptance, and route-following behavior. In scenarios with shared lanes and high demand, queue spillback and blocking effects naturally emerge due to limited downstream capacity and finite link storage, as explicitly modeled in SUMO. As a result, vehicles are often unable to perform discretionary lane changes, and lane utilization becomes constrained by queue formation and downstream congestion.

It is important to note that the model does not explicitly capture aggressive or non-compliant driving behavior that may occur in extreme congestion. However, the adopted modeling framework provides a realistic and sufficiently accurate representation of lane discipline under oversaturated conditions, particularly in terms of queue formation, spillback propagation, and restricted maneuverability at signalized intersections.

Queue spillback is explicitly captured in the simulation through finite link storage and downstream capacity constraints, which limit the inflow of vehicles when queues reach upstream intersections. As a result, vehicles may be blocked from entering an intersection even during green phases, reflecting realistic spillback conditions in urban networks.

The proposed HYB framework accounts for spillback effects through its state representation, which includes queue lengths and predicted short-term traffic conditions. This enables the controller to implicitly avoid allocating green time to blocked approaches and to prioritize movements with available downstream capacity, thereby reducing spillback propagation and maintaining stable operation under congested conditions.

All simulations were executed with a simulation step of 1 s. An initial stabilization (warm-up) period of 600 s was applied, during which results were excluded from statistical analysis. Each scenario was repeated 20 times with different random initial conditions. The results are reported as mean $\pm$ standard deviation (Std). These repeated runs were used as the basis for statistical comparison across different control strategies and performance indicators.

The proposed HYB framework was compared with fixed-time control (FT), actuated control (ACT), MP, and a standalone RL-based control strategy. For a fair and consistent comparison, the standalone RL baseline was implemented using the same PPO algorithm as employed in the system, but without the integration of the prediction, metaheuristic optimization, and surrogate evaluation modules. In this configuration, the PPO agent operates directly on the current system state and performs adaptive TSC based solely on real-time observations. This design ensures that the observed performance improvements of the proposed HYB framework are not due to differences in the underlying RL architecture, but rather result from the integration of additional modules that enhance decision-making capabilities.

The ACT baseline is implemented according to standard traffic engineering practice, using detector-based phase extension logic. For the implementation of the ACT control strategy, virtual loop detectors are placed on each incoming lane at a distance of approximately 30–50 m upstream of the stop line. These detectors are used to measure vehicle presence and queue formation, enabling phase extension and gap-out logic consistent with standard actuated control principles.

Each phase is associated with minimum and maximum green times ($g_{min} = 10$ s, $g_{max} = 60$ s), while phase termination is governed by gap-out and max-out conditions. A gap-out threshold of 2.5 s is used to determine the absence of demand. Intergreen times include a yellow interval of 3 s and an all-red interval of 1 s. The controller extends the green phase as long as vehicle presence is detected within the predefined time gap and terminates it when no demand is detected or when the maximum green time is reached. This control logic reflects typical real-world ACT strategies used in urban intersections.

The MP controller is implemented following the classical formulation, where at each decision step the phase with the highest pressure is selected. The pressure is computed as the difference between upstream and downstream queue lengths for each movement, and the phase with the highest total network pressure is selected. In scenarios where multiple phases yield identical pressure values, a fixed phase order is applied as a tie-breaking rule.

All baseline controllers operate under the same signal timing constraints and input data conditions.

The following key performance indicators were evaluated: average vehicle delay, average queue length, EV and PT time-to-pass, and control stability indicators (phase change frequency and average vehicle delay variance). It is important to note that, although pedestrian demand is included in the scenario design, the evaluation is limited to vehicle-oriented key performance indicators. Consequently, the impact of the proposed control strategy on pedestrian service levels is not explicitly quantified in the present analysis.

The experiments were performed on a workstation equipped with an 8-core/16-thread CPU, 32 GB of RAM, and an NVMe SSD, supported by a GPU used for the deep learning tasks. This configuration enabled parallel execution of SUMO simulations and stable evaluation of the proposed HYB framework.

To assess the real-time applicability of the proposed HYB framework, the computational requirements of the online control cycle were analyzed. The execution time includes state update, short-term traffic prediction, surrogate evaluation, and RL-based action selection.

The results indicate that the average computation time per control cycle is approximately 160 ms on the specified hardware platform, well below 0.25 s and consistent with values reported in recent real-time TSC studies. This satisfies typical operational requirements, given that TSC decisions are executed at discrete intervals on the order of several seconds. Therefore, the computational latency remains well within acceptable limits for real-time intersection-level deployment. The offline components (metaheuristic optimization and model training) are not executed within the control loop and therefore do not affect real-time performance.

All simulations and model training procedures were performed using fixed random seeds to ensure reproducibility. Additional simulation files and configuration details are provided in the Data Availability Statement. All experiments were implemented using Python (version 3.10) and SUMO, with standard ML libraries.

4.1. Overall System Performance Under Different Traffic Load Levels

Aggregated results of the basic system performance indicators for the FT, ACT, MP, RL, and the proposed HYB framework across all considered traffic demand levels are presented in Table 5.

Table 5. Average performance indicators under different traffic demand levels.

Demand Grade	Control Strategy	Average Vehicle Delay (s/veh)	Average Queue (veh)
G1	FT	18.6 ± 2.1	4.2 ± 0.8
	ACT	15.3 ± 1.8	3.5 ± 0.7
	MP	13.9 ± 1.6	3.1 ± 0.6
	RL	13.4 ± 1.7	3.0 ± 0.6
	HYB	12.1 ± 1.4	2.6 ± 0.5
G2	FT	34.8 ± 3.6	9.8 ± 1.7
	ACT	29.1 ± 3.2	8.4 ± 1.5
	MP	26.5 ± 2.9	7.6 ± 1.3
	RL	25.4 ± 3.1	7.5 ± 1.4
	HYB	22.6 ± 2.5	6.6 ± 1.1
G3	FT	62.4 ± 7.1	22.8 ± 4.1
	ACT	55.6 ± 6.5	19.9 ± 3.6
	MP	49.2 ± 5.9	17.4 ± 3.1
	RL	47.8 ± 6.8	16.9 ± 3.8
	HYB	41.3 ± 5.2	14.6 ± 2.7
G4	FT	116.8 ± 16.9	50.1 ± 9.8
	ACT	108.9 ± 16.1	46.0 ± 9.2
	MP	97.6 ± 14.8	41.0 ± 8.4
	RL	95.9 ± 17.2	39.8 ± 10.6
	HYB	80.9 ± 12.1	31.6 ± 7.2

The results show that increasing demand (G1–G4) leads to a marked increase in average vehicle delay and queue length. However, the proposed HYB framework consistently outperforms the reference control strategies at every demand level, maintaining the lowest average vehicle delays and the shortest average queues.

These trends are further illustrated in Figure 3. The boxplot distributions reveal that the proposed HYB framework not only maintains lower median average vehicle delays but also exhibits lower dispersion compared to the standard RL control strategy. Furthermore, the proposed HYB framework's competitive advantage becomes increasingly pronounced under high and oversaturated traffic conditions.

Across all considered scenarios, the proposed HYB framework achieves the lowest average vehicle delay medians and the smallest interquartile range compared with the conventional FT, ACT, MP, and RL control strategies.

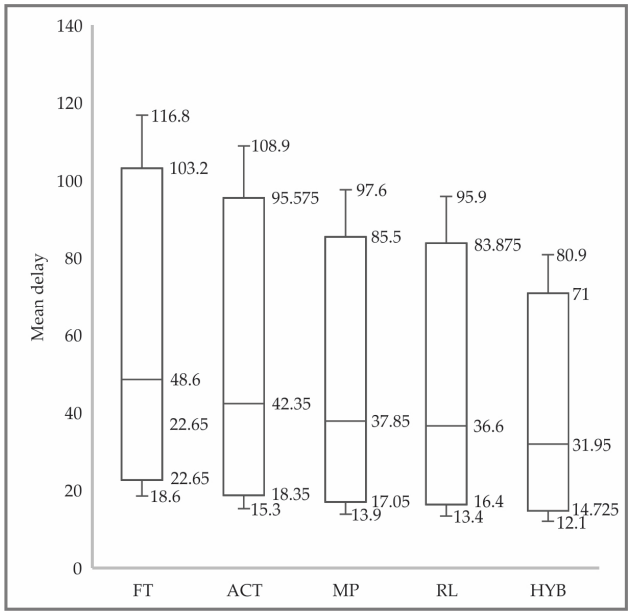


Figure 3. Distribution of average vehicle delay across 20 simulation runs, presented as boxplots (median, interquartile range, and variability).

4.2. Queue Dynamics Under Oversaturated Conditions

To provide a more detailed insight into system behavior under extreme conditions, Figure 4 illustrates the temporal evolution of the queue length on the critical approach for the G4–S3 scenario.

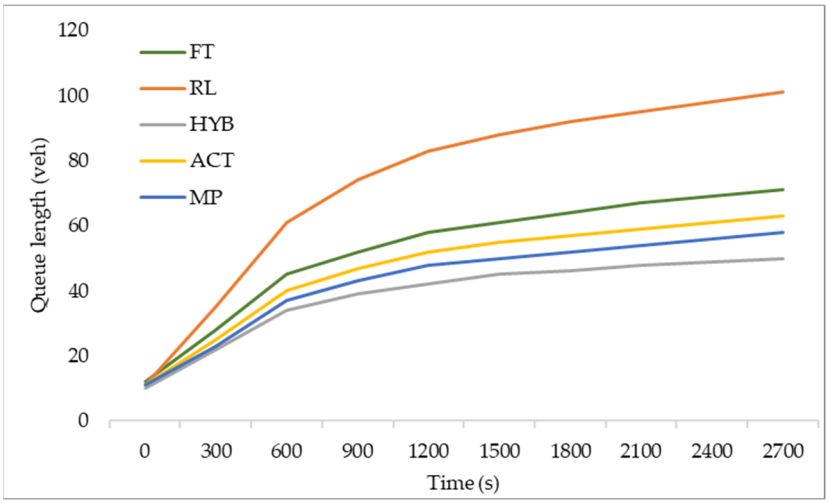


Figure 4. Queue length time series (critical approach).

The results indicate that:

- FT leads to continuous queue growth due to the non-adaptive nature of the signal plan and its inability to respond to fluctuations in traffic demand;
- ACT shows improved performance compared with FT by responding to vehicle detections; however, its reactive nature still results in moderate queue growth under higher traffic demand;
- MP controller achieves more efficient queue dissipation by prioritizing movements with the highest queue imbalance, resulting in shorter queues than FT and ACT but still exhibiting noticeable fluctuations in heavily saturated conditions;

- RL control strategy exhibits pronounced oscillations, with abrupt increases and decreases in queue length caused by frequent phase switching and reactive control behavior;
- The proposed HYB framework maintains the shortest and most stable queues throughout the simulation period, demonstrating its ability to balance responsiveness and stability.

The superior performance of the proposed HYB framework can be attributed to the integration of predictive and optimization modules, which reduce uncertainty in the control process and constrain the RL action space. In contrast, the standalone RL control strategy exhibits higher variability due to unconstrained exploration, leading to frequent phase switching and oscillatory behavior.

These findings confirm that the integration of prediction, optimization, and stabilization mechanisms enables more effective and stable control under oversaturated conditions.

4.3. EV and PT Priority Performance

Priority-related system performance is evaluated across the considered traffic demand levels (G1–G4) and priority scenarios (S2–S3). The results of the priority performance analysis are summarized in Table 6, where EV time-to-pass and PT delay reduction are presented. The reduction in PT delay is calculated relative to the FT control strategy, which is used as a baseline. All performance indicators are reported using the same definitions and units as in the main results tables, ensuring consistency and comparability.

Table 6. Priority-related performance indicators under priority scenarios (S2–S3).

Demand Grade	Control Strategy	EV Time-to-Pass (s)	PT Delay Reduction (%)
G1	ACT	32.8 ± 4.9	7.6
	MP	29.4 ± 4.3	10.1
	RL	26.8 ± 4.1	12.7
	HYB	24.7 ± 4.6	15.8
G2	ACT	43.2 ± 6.9	10.2
	MP	39.6 ± 6.1	13.4
	RL	36.9 ± 5.8	16.2
	HYB	34.2 ± 6.3	19.3
G3	ACT	55.7 ± 8.8	14.6
	MP	50.4 ± 7.9	18.7
	RL	47.1 ± 7.2	21.4
	HYB	43.7 ± 7.8	23.6
G4	ACT	85.6 ± 14.1	17.5
	MP	79.8 ± 12.6	21.9
	RL	74.2 ± 11.4	25.1
	HYB	69.8 ± 13.6	27.9

EV time-to-pass increases with rising traffic demand due to increased traffic conflicts and reduced operational flexibility under higher saturation levels. Across all demand grades, however, the proposed HYB framework consistently achieves the lowest EV time-to-pass compared to ACT, MP, and RL control strategies, indicating its effectiveness in prioritizing EV.

Furthermore, the effectiveness of PT priority becomes more pronounced under high and oversaturated traffic conditions (G3 and G4). The proposed HYB framework achieves the highest reduction in PT delay, indicating improved performance in handling priority vehicles under congested traffic conditions. Spillback-related effects are reflected in the observed queue lengths and average vehicle delay values, which serve as indirect indicators of blocking and congestion propagation. Lower queue lengths and reduced average vehicle delay under the proposed HYB framework indicate improved handling of spillback conditions compared to baseline strategies.

These trends are further illustrated in Figure 5, which presents EV time-to-pass across different traffic demand levels and operational scenarios (S1–S3). As shown in the figure, all control strategies exhibit increasing EV time-to-pass with rising traffic demand. However, the proposed HYB framework consistently outperforms ACT, MP, and standalone RL control strategies, and demonstrates clear improvements compared to the FT baseline. For clarity, scenario S1 is included as a baseline without priority, enabling a direct comparison with priority scenarios S2 and S3.

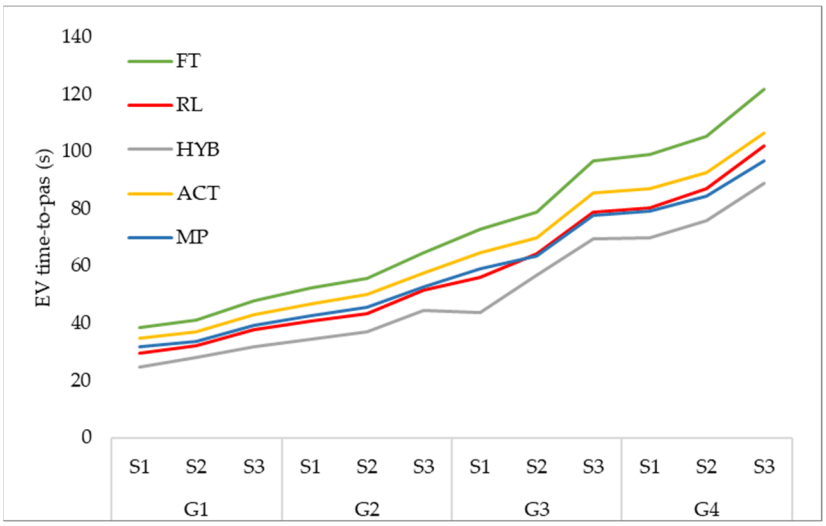


Figure 5. EV time-to-pass under scenarios S1–S3. Scenario S1 represents the baseline scenario without priority, while S2 and S3 correspond to priority scenarios.

In addition, the differences between control strategies are relatively smaller under low traffic demand (G1), while under higher demand levels (G3 and G4), the advantage of the proposed HYB framework becomes more pronounced. This behavior indicates that the system is particularly effective under congested conditions, where efficient priority handling is critical.

4.4. Analysis of Average Vehicle Delay

The percentage improvement in average vehicle delay achieved by the proposed HYB framework compared with the control strategies for different traffic demand levels (G1–G4) is illustrated in Figure 6. This improvement is computed as the relative reduction in average vehicle delay with respect to the corresponding reference TSC, according to (16):

$$Improvement = \frac{x_{baseline} - x_{HYB}}{x_{baseline}} \cdot 100[\%] \quad (16)$$

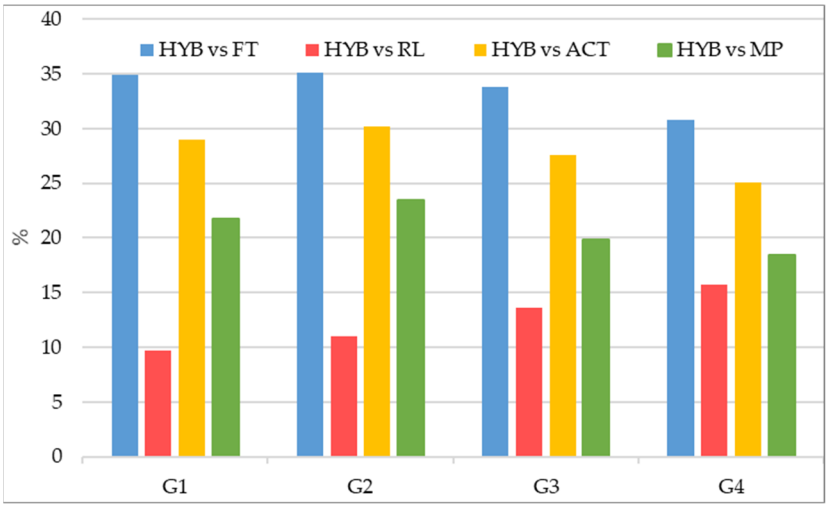


Figure 6. Percentage improvement in average vehicle delay achieved by the proposed HYB framework relative to FT and RL under different demand grades.

The results show that the proposed HYB framework achieves a significant reduction in average vehicle delay compared with FT signal control across all analyzed scenarios, with improvements ranging from 30.8% to 35.1%. The largest relative benefit is observed at moderate traffic demand levels (G2), indicating that the proposed HYB framework effectively balances adaptability and stability under variable traffic flow conditions.

Compared with the ACT and MP control strategies, the proposed HYB framework also demonstrates consistent improvements across all demand levels. The reduction in average vehicle delay relative to ACT ranges from approximately 25.1% to 30.2%, while the improvement compared with MP varies between 18.4% and 23.5%. These results indicate that although ACT and MP controllers provide better performance than conventional FT control, the proposed HYB framework achieves superior efficiency by combining prediction, optimization, and adaptive learning mechanisms.

When compared with the RL-based control strategy, the proposed HYB framework demonstrates increasing relative improvement as traffic demand rises, from 9.7% in level G1 to 15.7% in level G4. This trend suggests that purely RL-based control becomes less stable under high congestion levels, while the proposed HYB framework maintains more consistent performance and lower average vehicle delay variability.

In addition to general traffic performance, a focused analysis was conducted for priority vehicles. The results indicate that the proposed HYB framework achieves a consistent reduction in EV time-to-pass and PT delay compared to all baseline control strategies.

Compared to the RL baseline, the proposed HYB framework reduces EV travel time through the intersection by approximately 12–14%, while PT delay is reduced by approximately 8–10%. In contrast, conventional strategies (FT and ACT) show limited capability in handling priority requests, while the MP control strategy provides partial improvements but lacks explicit priority mechanisms. These results confirm that the integration of priority-aware control mechanisms within the proposed HYB framework enables effective prioritization of critical traffic flows without compromising overall system stability.

Statistical analysis was conducted using the nonparametric Mann–Whitney U test to compare the performance of the proposed HYB framework against the RL baseline, which serves as the primary reference for statistical evaluation. All tests were performed as two-sided, and statistical significance was evaluated at a level of $p < 0.05$. To account for multiple pairwise comparisons, the Bonferroni correction was applied. The statistical testing procedure is described in Section 3.10.

The analysis was performed for multiple key performance indicators, including average vehicle delay, queue length, EV time-to-pass, PT delay reduction, and phase-switch frequency. In all scenarios, the proposed HYB framework demonstrated statistically significant improvements compared to the RL baseline, with corrected $p < 0.05$. In addition, effect sizes were computed using rank-biserial correlation, indicating consistently large effects across all evaluated metrics. Detailed statistical results, including p-values and effect sizes for all performance indicators, are presented in Table 7.

Table 7. Statistical test results for pairwise comparison between the proposed HYB framework and the RL baseline across KPI.

KPI	Comparison	p-Value	Corrected p	Effect Size
Average vehicle delay	HYB vs. RL	<0.001	<0.001	0.72
Queue length	HYB vs. RL	<0.001	<0.001	0.69
EV delay	HYB vs. RL	<0.001	<0.001	0.75
PT delay	HYB vs. RL	<0.001	<0.001	0.7

These results confirm that the performance improvements achieved by the proposed HYB framework are not due to random variability, but represent a consistent and statistically robust advantage. While Figure 6 reports percentage improvements relative to multiple benchmarks, statistical testing is conducted specifically against the RL baseline. This baseline serves as the primary reference for evaluating the efficacy of the proposed adaptive control.

4.5. Control Stability Analysis

Control stability is summarized in Table 8, where the ACT, MP, RL, and HYB control strategies are compared in terms of phase change frequency and average vehicle delay variance. The delay variance is computed as the variance of average vehicle delay across 20 repeated simulation runs for a given control strategy and demand scenario. This metric reflects the temporal stability and consistency of control performance under stochastic traffic conditions.

Table 8. Control stability indicators under high and oversaturated traffic demand (aggregated across scenarios G1–G4).

Control Strategy	Phase Switches (/h)	Delay Variance
ACT	95 ± 18	410 ± 70
MP	128 ± 24	455 ± 76
RL	148 ± 29	520 ± 82
HYB	72 ± 12	320 ± 50

Delay variance is calculated as the variance of average vehicle delay across 20 simulation runs. The results show that:

- RL control strategy results in a high number of phase changes per hour, indicating aggressive and oscillatory behavior caused by continuous adaptation to short-term traffic fluctuations;
- ACT demonstrates a more stable behavior, with fewer phase changes than RL due to predefined minimum green and gap-out conditions. However, because it reacts only to local detector inputs, the variability of average vehicle delay remains relatively high;

- MP controller exhibits dynamic phase switching as it continuously selects movements with the highest queue imbalance. This leads to improved queue dissipation but may also increase control variability under fluctuating traffic demand;
- In contrast, the proposed HYB framework significantly reduces the number of phase changes while simultaneously achieving lower average vehicle delay and delay variance, indicating improved control stability and a more balanced response to changing traffic conditions.

This trade-off between efficiency and stability is clearly illustrated in Figure 7, which confirms that the proposed HYB framework represents the dominant solution.

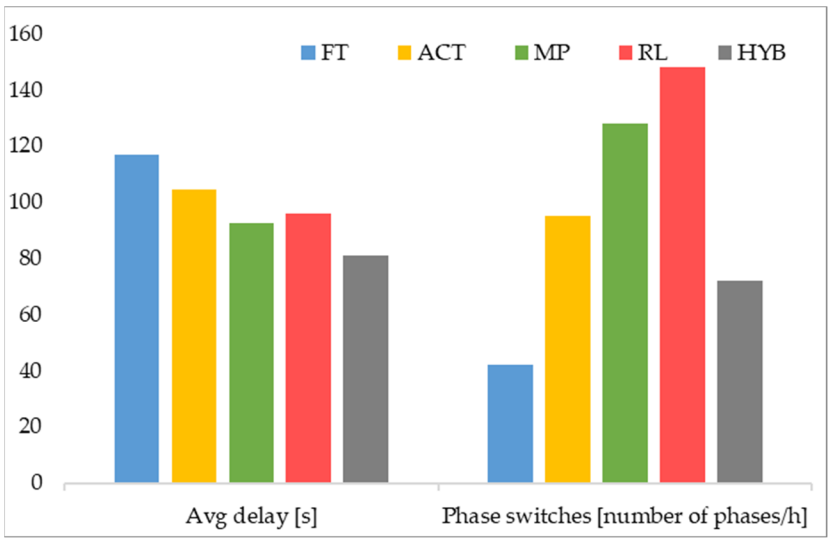


Figure 7. Pareto-based comparison of control strategies in terms of average vehicle delay and stability.

The Pareto analysis highlights the trade-off between average vehicle delay and control stability, represented by phase-switching frequency and delay variance. The FT strategy exhibits the lowest phase-switching frequency but results in the highest delay, whereas the RL control strategy reduces delay at the cost of excessive phase switching and higher variability. Although the ACT and MP controllers offer intermediate performance, the proposed HYB framework achieves the lowest overall delay with moderate phase changes and reduced variability. Consequently, it provides the most balanced solution, optimizing both efficiency and control stability.

4.6. Robustness and Sensitivity Analysis

To further evaluate the reliability of the proposed HYB framework, an additional robustness analysis was conducted under non-ideal operating conditions. Two types of disturbances were considered: measurement uncertainty in traffic detector data and sudden increases in traffic demand. The first test introduced stochastic noise into the input traffic flow and queue measurements to simulate sensor inaccuracies. The second test evaluated system behavior under escalated traffic demand levels relative to the calibrated baseline scenario.

Measurement noise was modeled as additive stochastic perturbation applied to traffic flow and queue length inputs. Specifically, zero-mean Gaussian noise was introduced, with Std proportional to the nominal value ($\sigma = 5\%$, 10%, and 15%), in order to simulate detector inaccuracies.

Traffic demand perturbations were implemented by scaling the arrival rates (λ) of the calibrated Poisson processes. Demand levels were increased by 10%, 20%, and 30%

relative to the baseline scenario, while preserving the original turning ratios across all control strategies.

In addition to measurement noise and demand variability, the impact of input data latency was also considered, as real-world detector data are subject to delays caused by sampling frequency, communication latency, and processing time. To account for this, latency effects were approximated through controlled perturbations of input data used in the prediction and control modules, thereby reflecting the impact of outdated or imperfect state information on decision-making.

Notably, higher latencies (e.g., 30–60 s), typically associated with aggregated data from legacy loop detectors, were not explicitly modeled. This is because the proposed HYB framework is designed for near real-time data streams, leveraging the capabilities of modern sensing and communication technologies. The results indicate that system performance gradually degrades under increased input uncertainty. However, the system maintains stable operation with only moderate performance degradation, confirming its robustness under realistic conditions where perfect real-time data are not available.

The results suggest that, as a preliminary observation, the proposed framework remains operational and maintains stable performance for latencies up to approximately 10 s; however, this finding is based on qualitative assessment, and higher delays appear to result in a noticeable degradation in performance.

Overall, the proposed HYB framework preserves stable performance under both types of disturbances. Although the average vehicle delay and queue length increase with higher noise levels and demand growth, the degradation remains moderate compared with the RL baseline. In particular, the system shows lower sensitivity to input perturbations, which can be attributed to the combined effect of prediction, surrogate evaluation, and constrained RL-based fine-tuning.

Under detector noise, the proposed HYB framework maintains relatively low delay variability and avoids excessive phase switching, indicating that the control policy is not overly dependent on exact state measurements. Similarly, under increased traffic demand, the proposed HYB framework continues to outperform the RL baseline, especially in terms of average vehicle delay and queue stability. These findings confirm that the system is robust for application in realistic traffic environments characterized by uncertainty and fluctuating conditions.

The quantitative results of the robustness analysis are summarized in Table 9.

Table 9. Robustness analysis under detector noise and traffic demand variation.

Disturbance Level	RL Average Vehicle Delay (s/veh)	HYB Average Vehicle Delay (s/veh)	RL Phase Switches (/h)	HYB Phase Switches (/h)
Baseline (0% noise and demand increase)	95.9	80.9	148	72
5% detector noise	99.6	83.4	151	75
10% detector noise	104.8	86.7	156	79
15% detector noise	111.2	91.5	163	84
+10% traffic demand	107.4	89.8	154	77
+20% traffic demand	118.9	97.6	161	82
+30% traffic demand	132.7	108.8	170	89

The robustness results show that both detector noise and traffic demand growth lead to performance degradation for both considered control strategies (RL and HYB). However, the proposed HYB framework exhibits a slower increase in average vehicle delay and a lower rise in phase-switching frequency compared with the RL control strategy. This

confirms that the system is less sensitive to input uncertainty and better suited for stable operation under disturbed traffic conditions.

4.7. Contribution of the Proposed HYB Framework and Computational Feasibility

To evaluate the contribution of individual components of the proposed HYB framework, an ablation analysis was conducted for the representative high-demand priority scenario G3–S2, with the quantitative results summarized in Table 10. The standalone RL-based control module exhibits the highest delay, queue length, and instability, reflected in a high number of phase switches and increased delay variance.

Table 10. Ablation study results for the representative high-demand priority scenario G3–S2 (mean ± Std over 20 independent runs).

Model Configuration	Average Vehicle Delay (s)	Queue Length [1] (veh)	EV Time (s)	PT Delay Reduction [2] (%)	Phase Switches (/h)	Delay Variance
RL-only controller	45.6 ± 5.9	15.8 ± 3.1	45.4 ± 6.8	20.8	142 ± 25	486 ± 71
Prediction-assisted RL	43.9 ± 5.4	15.1 ± 2.9	43.8 ± 6.4	22.1	127 ± 21	431 ± 63
Optimization-assisted RL	42.1 ± 5.0	14.4 ± 2.6	42.5 ± 6.1	22.9	101 ± 17	372 ± 56
Full HYB framework	40.4 ± 4.7	13.7 ± 2.4	41.2 ± 5.9	24.1	78 ± 13	309 ± 47

[1] Queue length is expressed in number of vehicles, consistent with the main results section. [2] PT delay reduction (%) is calculated relative to the FT control strategy.

These results provide direct and quantitative evidence of the computational feasibility of the proposed HYB framework. The inclusion of the prediction module improves overall efficiency and priority handling, as reflected in reduced average vehicle delay and improved EV/PT performance, due to the anticipatory capability introduced by short-term traffic forecasting. Further improvements are achieved by incorporating the optimization module, which constrains the action space to high-quality signal plans, resulting in more stable control behavior and reduced phase switching.

The proposed HYB framework achieves the best performance across all indicators, achieving the lowest delay, shortest queues, improved priority performance, and the most stable control. These results indicate that the observed improvements arise from the complementary integration of prediction, optimization, and adaptive control, rather than from any single component.

From the computational perspective, the proposed framework separates offline and online processes. Computationally intensive tasks, such as model training and signal plan generation, are performed offline, while online operation is limited to lightweight inference and decision-making. The control logic is executed at each simulation step based on the current traffic state and predicted short-term conditions. The obtained results indicate that the control cycle operates within 160 ± 22 ms, supporting the applicability of the proposed HYB framework in real-time intersection control.

To provide a more detailed assessment of the real-time performance, the computational time of the main online modules was measured, with the results presented in Table 11. The reported values represent the average and standard deviation over 20 independent simulation runs.

Table 11. Average computational time of main online modules (per control cycle).

Module	Mean Time (ms)	Std (ms)	Max (ms)
State update (TraCI)	45	8	65
Prediction inference (DL)	62	10	85
Surrogate evaluation (XGBoost)	18	4	30
PPO action selection	35	6	50
Total per control cycle	160	22	210

These findings demonstrate that the proposed HYB framework satisfies real-time operational requirements, with execution times consistently well below 0.25 s per control cycle. As shown in Table 11, all individual modules exhibit low execution times, with no single component exceeding 70 ms on average, indicating the absence of critical computational bottlenecks. The maximum observed execution time per control cycle remains below 250 ms, confirming robust performance even under peak computational load.

Although the reported timings correspond to the specified hardware configuration, the computational structure of the system is dominated by lightweight inference operations. This enables deployment on less powerful edge devices, particularly due to the use of a surrogate model that significantly reduces computational overhead compared to simulation-based evaluation.

The computational complexity scales approximately linearly with the number of controlled intersections and input features, as the main online components operate independently per control unit, supporting scalability to larger traffic networks.

All timing measurements were obtained using Python time profiling during simulation execution on the hardware configuration described in Table 3. These results confirm that the proposed HYB framework meets real-time operational requirements and is suitable for deployment in practical TSC systems.

5. Discussion

An analysis of the obtained results indicates that the differences among the considered TSC strategies increase significantly with rising traffic flow saturation levels. Under unsaturated conditions (G1), where sufficient capacity reserves are available, the improvements achieved by advanced control strategy are relatively moderate, as conventional FT control can still provide an acceptable level of service. However, as the system transitions into highly saturated and oversaturated regimes (G3 and G4), the advantages of the proposed HYB framework become dominant, clearly demonstrating its robustness and ability to operate stably under demanding traffic conditions.

In contrast to the standalone RL-based control, which rely primarily on reacting to the current system state, the proposed HYB framework integrates traffic flow prediction, metaheuristic optimization, and adaptive learning. This finding is consistent with recent studies indicating that hybrid architectures combining prediction and RL provide improved robustness compared with standalone RL-based control. This multilayer architecture enables anticipatory decision-making, thereby reducing the impact of short-term fluctuations and stochastic disturbances. As a result, lower average vehicle delays, shorter queue lengths, and reduced performance variability are achieved, which represent key indicators of control stability. Integrating a surrogate model before the RL stage significantly enhances the stability and convergence speed of the PPO agent. By guiding action space exploration toward high-quality signal plans from the outset, this control strategy reduces reward variance and mitigates the risk of unstable learning during the early training phases. The observed improvements under high and oversaturated traffic conditions are also influenced by the underlying traffic behavior modeling. In particular, lane-changing dynamics in the simula-

tion are governed by the LC2013 model, which ensures that vehicle interactions remain constrained by safety conditions and available gaps. As traffic demand increases, more frequent lane-change attempts occur; however, these maneuvers remain physically feasible and do not lead to unrealistic traffic patterns. This contributes to a realistic representation of congestion dynamics and supports the validity of the obtained results.

A particular aspect of the discussion concerns the prioritization of EV and PT. The results demonstrate that the proposed HYB framework enables reliable priority realization even under conditions of intense traffic flow conflicts. Unlike aggressive adaptive strategies, where improvements for priority vehicles are often achieved at the expense of system stability (e.g., through an excessive number of phase changes), the proposed HYB framework successfully preserves a balance between priority handling and overall control stability.

Furthermore, the Pareto stability analysis indicates that the proposed HYB framework achieves a favorable compromise between efficiency and stability, yielding low average vehicle delay with a moderate number of signal phase changes. This finding is particularly important from the perspective of real-world implementation, as excessive phase switching may negatively affect equipment reliability, traffic safety, and overall system acceptability in practice.

Although this analysis focused on a single isolated intersection, the results strongly suggest that the proposed HYB framework is well-suited for real-world urban environments. It is particularly effective for areas characterized by high traffic demand variability and the frequent presence of priority vehicles.

6. Conclusions

This paper introduces the proposed HYB framework that integrates traffic flow prediction, metaheuristic optimization, and RL with the aim of improving both the efficiency and stability of signalized intersection operation. Simulation-based evaluation conducted under different traffic demand levels demonstrates that the system achieves a significant reduction in average vehicle delay compared with conventional and adaptive control strategies, including FT, ACT, MP and RL. The improvement reaches up to 35% relative to FT control and remains consistently superior compared with ACT, MP, and RL control strategies across all considered demand levels.

The results confirm that the advantages of the proposed HYB framework are particularly pronounced under high and oversaturated traffic conditions, where the requirements for adaptability and control stability are most critical. In addition to enhancing core performance metrics, the proposed HYB framework demonstrates a robust capacity for preserving system stability. This is evidenced by reduced performance variability and a highly favorable trade-off between average vehicle delay and signal switching frequency. Nevertheless, the authors acknowledge that real-world deployments may involve higher latency levels, which could further reduce the observed performance gains.

The analysis of EV and PT priority further highlights the advantages of the proposed HYB framework compared to conventional and adaptive control strategies. Specifically, the results confirm the practical value of the proposed HYB framework, as performance improvements are achieved without compromising overall system stability. This capability represents an important step toward the deployment of intelligent TSC systems in real urban environments.

Although the proposed HYB framework demonstrates strong performance for isolated intersection control, its extension to corridor-level and network-wide applications requires further investigation. In particular, additional mechanisms for coordination between

adjacent intersections, handling of spillback phenomena, and scalability of the state and action representations must be developed and validated.

Notably, the present study is based on a simulation analysis of a single isolated intersection, and therefore the obtained results should be interpreted within this context. While the use of a calibrated microsimulation environment enables controlled and reproducible evaluation of different control strategies, real-world traffic systems may involve additional complexities not fully captured in the simulation framework. Consequently, the reported performance improvements represent results under controlled conditions and should be considered as indicative rather than directly generalizable to large-scale urban networks.

While the proposed HYB framework demonstrates significant improvements in vehicular traffic performance, its impact on pedestrian service levels has not been explicitly evaluated. Future work will address this limitation by incorporating pedestrian-oriented metrics and analyzing multimodal traffic interactions. Furthermore, the framework will be extended to include environmental performance indicators, such as CO_2 emissions and fuel consumption, based on more detailed vehicle-level data and enhanced simulation capabilities.

Finally, future research will explore multi-agent coordination and large-scale simulations to evaluate scalability under sensor noise, communication latency, and corridor-level or network-wide conditions.

Author Contributions: Conceptualization, B.L., G.P. and D.M.; methodology, B.L. and G.P.; software, B.L.; validation, B.L., G.P. and D.M.; formal analysis, B.L. and G.P.; investigation, B.L., G.P. and Ž.Ć.; resources, G.P. and D.M.; data curation, B.L., G.P. and Ž.Ć.; writing—original draft preparation, B.L.; writing—review and editing, G.P., D.M., Ž.Ć. and S.D.; visualization, B.L. and G.P.; supervision, Ž.Ć., D.M. and S.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding specifically allocated to this study or its publication.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to (ongoing re-search activities related to a doctoral dissertation and the non-public availability of the implementation code).

Acknowledgments: This research was financially supported by the Ministry of Science, Technological Development, and Innovation of the Republic of Serbia (Contract No. 451-03-33/2026-03/200109).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ITS	Intelligent Transportation Systems
TSC	Traffic Signal Control
MP	Max-Pressure Control
ML	Machine Learning
DL	Deep Learning
LSTM	Long Short-Term Memory
GNNs	Graph Neural Networks
RL	Reinforcement Learning
DRL	Deep Reinforcement Learning
DQN	Deep Q-Network
PPO	Proximal Policy Optimization
A2C	Advantage Actor-Critic

PSO	Particle Swarm Optimization
GWO	Grey Wolf Optimizer
HYB	Hybrid
EV	Emergency Vehicle
PT	Public Transport Vehicles
SUMO	Simulation of Urban Mobility
V2I	Vehicle-to-Infrastructure
V2X	Vehicle-to-Everything
YOLO	You Only Look Once
TraCI	Traffic Control Interface
API	Application Programming Interface
ETA	Estimated Time of Arrival
GRU	Gated Recurrent Unit
MSE	Mean Squared Error
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
R^2	Coefficient of Determination
SPaT	Signal Phase and Timing
Std	Standard Deviation
G1	Low traffic demand
G2	Medium traffic demand
G3	High traffic demand
G4	Oversaturated traffic regime
S1	Scenario with no priority granted to specific vehicle groups
S2	Scenario with priority for emergency vehicles
S3	Scenario with priority for public transport vehicles
S4	Mixed scenario (emergency vehicles, public transport, and increased pedestrian flows)
FT	Fixed-Time Control
ACT	Actuated Control

References

1. Papageorgiou, M.; Diakaki, C.; Dinopoulou, V.; Kotsialos, A.; Wang, Y. Review of road traffic control strategies. *Proc. IEEE* **2003**, *91*, 2043–2067. [CrossRef]
2. Hunt, P.B.; Robertson, D.I.; Bretherton, R.D.; Royle, M.C. The SCOOT on-line traffic signal optimisation technique. *Traffic Eng. Control* **1982**, *23*, 190–192.
3. Lowrie, P.R. *Scats, Sydney Co-Ordinated Adaptive Traffic System: A Traffic Responsive Method of Controlling Urban Traffic*; Roads and Traffic Authority: Darlinghurst, NSW, Australia, 1990.
4. Mirchandani, P.; Head, L. A real-time traffic signal control system: Architecture, algorithms, and analysis. *Transp. Res. Part C Emerg. Technol.* **2001**, *9*, 415–432. [CrossRef]
5. Maadi, S.; Stein, S.; Hong, J.; Murray-Smith, R. Real-time adaptive traffic signal control in a connected and automated vehicle environment: Optimisation of signal planning with reinforcement learning under vehicle speed guidance. *Sensors* **2022**, *22*, 7501. [CrossRef]
6. Varaiya, P. Max pressure control of a network of signalized intersections. *Transp. Res. Part C Emerg. Technol.* **2013**, *36*, 177–195. [CrossRef]
7. Ma, X.; Tao, Z.; Wang, Y.; Yu, H.; Wang, Y. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transp. Res. Part C Emerg. Technol.* **2015**, *54*, 187–197. [CrossRef]
8. Wei, H.; Zheng, G.; Yao, H.; Li, Z. IntelliLight: A reinforcement learning approach for intelligent traffic light control. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018; pp. 2496–2505. [CrossRef]
9. Li, L.; Lv, Y.; Wang, F.Y. Traffic signal timing via deep reinforcement learning. *IEEE/CAA J. Autom. Sin.* **2016**, *3*, 247–254. [CrossRef]
10. Nikolić, F.; Čanađija, M. Machine learning of structure–property relationships: An application to heat generation during plastic deformation. *Facta Univ. Ser. Mech. Eng.* **2025**, *23*, 687–707. [CrossRef]
11. Li, Y.; Yu, R.; Shahabi, C.; Liu, Y. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv* **2017**, arXiv:1707.01926. [CrossRef]

12. Zhang, J.; Zheng, Y.; Qi, D. Deep spatio-temporal residual networks for citywide crowd flows prediction. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), San Francisco, CA, USA, 4–9 February 2017; pp. 1655–1661. [CrossRef]
13. Yu, B.; Yin, H.; Zhu, Z. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. In Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI-ECAI 2018), Stockholm, Sweden, 13–19 July 2018; pp. 3634–3640. [CrossRef]
14. Chen, C.; Wei, H.; Xu, N.; Zheng, G.; Yang, M.; Xiong, Y.; Xu, K.; Li, Z. Toward a thousand lights: Decentralized deep reinforcement learning for large-scale traffic signal control. *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 3414–3421. [CrossRef]
15. Prashanth, L.A.; Bhatnagar, S. Reinforcement learning with function approximation for traffic signal control. *IEEE Trans. Intell. Transp. Syst.* **2010**, *12*, 412–421. [CrossRef]
16. El-Tantawy, S.; Abdulhai, B.; Abdelgawad, H. Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (MARLIN-ATSC): Methodology and large-scale application on downtown Toronto. *IEEE Trans. Intell. Transp. Syst.* **2013**, *14*, 1140–1150. [CrossRef]
17. Liang, X.; Du, X.; Wang, G.; Han, Z. A deep reinforcement learning network for traffic light cycle control. *IEEE Trans. Veh. Technol.* **2019**, *68*, 1243–1253. [CrossRef]
18. Olayode, I.O.; Tartibu, L.K.; Okwu, M.O.; Ukaegbu, U.F. Development of a hybrid artificial neural network-particle swarm optimization model for the modelling of traffic flow of vehicles at signalized road intersections. *Appl. Sci.* **2021**, *11*, 8387. [CrossRef]
19. Gokulachandran, J.; Thenarasu, M.; Pothkanoori, B.; Narassima, M.S.; Tirkolaee, E.B. Reliability prediction and process parameter optimization of welded joints: Artificial neural network and fuzzy logic. *Facta Univ. Ser. Mech. Eng.* **2025**, *23*, 945–969. [CrossRef]
20. Chu, T.; Wang, J.; Codecà, L.; Li, Z. Multi-agent deep reinforcement learning for large-scale traffic signal control. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 1086–1095. [CrossRef]
21. Kolat, M.; Kővári, B.; Bécsi, T.; Aradi, S. Multi-agent reinforcement learning for traffic signal control: A cooperative approach. *Sustainability* **2023**, *15*, 3479. [CrossRef]
22. Zhang, Y.; Wang, S.; Jia, D.; Fan, P.; Jiang, R.; Gu, H.; Chow, A.H.F. Toward dependency dynamics in multi-agent reinforcement learning for traffic signal control. *arXiv* **2025**, arXiv:2502.16608. [CrossRef]
23. Jia, W.; Ji, M. Multi-agent deep reinforcement learning for large-scale traffic signal control with spatio-temporal attention mechanism. *Appl. Sci.* **2025**, *15*, 8605. [CrossRef]
24. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95—International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; pp. 1942–1948. [CrossRef]
25. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]
26. Qiao, Z.; Ke, L.; Zhang, G.; Wang, X. Adaptive collaborative optimization of traffic network signal timing based on immune-fireworks algorithm and hierarchical strategy. *Appl. Intell.* **2021**, *51*, 6951–6967. [CrossRef]
27. Kouidri, C.; Bouiadjra, R.B.; Mahi, F. ACO-DSOS hybrid approach to enhance traffic signal optimization. *Int. J. Transp. Dev. Integr.* **2024**, *8*, 9–17. [CrossRef]
28. Li, X.; Guo, S.; Sun, D.; Cao, L.; Li, C.; Tian, S.; Peng, L.; Qi, Y. A rolling bearing fault diagnosis method based on extreme learning machine optimized by improved whale optimization algorithm. *Facta Univ. Ser. Mech. Eng.* **2025**, *in press*. [CrossRef]
29. Qiu, L.; Zhang, D.; Tian, Y.; Al-Nabhan, N. Deep learning-based algorithm for vehicle detection in intelligent transportation systems. *J. Supercomput.* **2021**, *77*, 11083–11098. [CrossRef]
30. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 7464–7475. [CrossRef]
31. Baghel, A.; Srivastava, A.; Tyagi, A.; Goel, S.; Nagrath, P. Analysis of Ex-YOLO algorithm with other real-time algorithms for emergency vehicle detection. In *Proceedings of the First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019), Chandigarh, India, 12–13 October 2019*; Springer: Singapore, 2020; pp. 607–618. [CrossRef]
32. Zhao, L.; Song, Y.; Zhang, C.; Liu, Y.; Wang, P.; Lin, T.; Li, H. T-GCN: A temporal graph convolutional network for traffic prediction. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 3848–3858. [CrossRef]
33. Mzili, T.; Mzili, I.; Riffi, M.E.; Pamucar, D.; Simic, V.; Abualigah, L.; Almohsen, B. Hybrid genetic and penguin search optimization algorithm (GA-PSEOA) for efficient flow shop scheduling solutions. *Facta Univ. Ser. Mech. Eng.* **2024**, *22*, 77–100. [CrossRef]
34. Haghshenas, S.S.; Seraji, M.H.M.; Haghshenas, S.S.; Guido, G.; Astarita, V.; Simic, V.; Marinkovic, D. Enhancing support vector machine performance in forecasting the number of vehicles involved in traffic crashes via metaheuristic optimization algorithms. *Facta Univ. Ser. Mech. Eng.* 2025, *in press*. [CrossRef]
35. Liang, Y.; Ren, Z.; Wang, L.; Liu, H.; Du, W. Surrogate-assisted cooperative signal optimization for large-scale traffic networks. *Knowl.-Based Syst.* **2021**, *234*, 107542. [CrossRef]

36. Zhang, H.; Sun, J.; Tian, Y. Accelerated safety testing for highly automated vehicles: Application and capability comparison of surrogate models. *IEEE Trans. Intell. Veh.* **2023**, *9*, 2409–2418. [CrossRef]
37. Jang, I.; Jang, Y.H. Prediction of contact distribution on rough surfaces using deep learning algorithms. *Facta Univ. Ser. Mech. Eng.* **2025**, *23*, 757–785. [CrossRef]
38. Lopez, P.A.; Behrisch, M.; Bieker-Walz, L.; Erdmann, J.; Flötteröd, Y.P.; Hilbrich, R.; Lucken, L.; Rummel, J.; Wagner, P.; Wießner, E. Microscopic traffic simulation using sumo. In *Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018*; IEEE: New York, NY, USA, 2018; pp. 2575–2582. [CrossRef]

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