Optimization of Traffic Signal Timing for Urban Intersections

Table of contents

Abstract

Effective traffic management represents an essential factor to smart urban systems and requires adaptive and accurate prediction models. This study gives a complete time series analysis of traffic volume data in 36 different locations. Each location is treated as a separate feature. Two input types of early traffic features were extracted and integrated: data given each 15 minutes and supervised machine learning algorithms Random Forest Regressor and Extreme Gradient Boosting (XGBoost). The performances of the models were evaluated using the coefficient of determination (R^2) and root mean squared error (RMSE). The results indicate that both models did well generally, with XGBoost performing slightly better, especially at locations traffic_5, traffic_22, and traffic_23, where it provided an R^2 of greater than 0.95 and RMSE of less than 0.03. However, at traffic_27, both met failure, which indicates possible anomalies unique to this location. The normalized dataset offers more stability and comparability to models. This scalable forecasting framework serves for actual traffic management and therefore gives a strong building block for intelligent transportation system and urban solutions.

1. INTRODUCTION

Urbanization and the massive influx of vehicular traffic into cities have created very good problems for traffic management, particularly at signalized intersections. These ones are very critical nodes in the transportation network being the major responsible for congestion, increased travel time, fuel wastage, and environmental pollution. On the one hand, conventional signal control systems-set in black and white through some pre-timed or actuated signal scheme modalities-do not distinguish well parameters that dynamically interact with a complexly changing traffic system. Novel intelligent and adaptive signal control techniques that will therefore work to maximize efficiency of flow, reduce congestion, and improve mobility in urban centres are thus demanded.

In the recent past, ML techniques are being proposed as useful instruments to untangle the processes underlying the optimization of traffic signal settings. In contrast to conventional methods, which mostly are reliant on static rules and assumptions, ML algorithms learn patterns from real-time and historical traffic data. These patterns can analyse large amounts of data to bring to light hidden correlations regarding the traffic flow characteristics and signal timing strategies. Therefore, the road towards an adaptive, responsive, and efficient traffic signal control system can be taken by feeding learning into this data-driven system.

ML optimization for traffic signal timing encompasses data gathering, feature engineering, model selection, training, validation, and deployment-commissioned works. The input comprises a wide range of data coming from sources such as loop detectors, cameras, GPS devices, and traffic management centres. Types of data collected may include traffic volumes, vehicle speeds, signal phase timings, queue lengths, and so on. This data provides enough information for machine learning algorithms to predict traffic behaviour and optimize signal timing plans to minimize delays and stops and reduce fuel consumption while maximizing throughput. Optimizing Traffic Signal Timing using Machine Learning involves various steps like data gathering from different sources such as loop detectors, cameras, GPS, etc., and finally from traffic management centres. Other data may include traffic volumes, vehicle speeds, queue lengths, and signal phase timings. Machine learning algorithms use this information to infer the prediction of optimum signal timing plans that minimize delay and stop expenditure as well as fuel consumption while maximizing throughput.

Several machine learning models have been examined in this area, supervised learning methods like decision trees, support vector machines, and neural networks, in addition to reinforcement learning techniques focusing on sequences of decisions and maximization of rewards. Out of all this, deep reinforcement learning is promisingly modelling the signal control problem as a Markov Decision Process (MDP), with the agent learning the optimal policy through interaction with the environment. It also involves a particular research and possible method of using such importation for developing smart rural towns.

The real-time merging of data with IoT technologies opens up avenues for more robust adaptive traffic control systems. Smart sensors and edge computing devices continuously monitor traffic conditions and feed that data into ML models that can be continuously optimized. This real-time adaptability goes beyond improving traffic conditions; it works well with others in emergency vehicle prioritization, pedestrian safety, and environmental sustainability.

Even though it has so much capability, do face challenges for implementation ML-based traffic signal optimization systems, like inadequate quality of data, generalization of models on unseen scenarios, computational challenges for real time, and the need to coordinate traffic authorities from other jurisdictions. However, machine learning is continuously changing dimensions intelligence transportation systems because of advancements from computation and availability of data and sophistication of algorithms.

This paper describes a comprehensive approach to the optimization of traffic signal timing at urban intersections by means of machine learning techniques. The work is directed towards building predictive and adaptive models that will be able to deal with dynamic changes in traffic patterns, thereby improving the flow of traffic and reducing congestion within urban settings.

1.1 BACKGROUND AND SCOPE

Urbanization has led to the rapid growth of cities and has increased the demand for efficient traffic management systems. Traffic signal systems at intersections are one of the most important elements of urban traffic management. In metropolitan cities where road networks are dense with high vehicle flow, timing of traffic signals inefficiently could account for high delays, congestion, and accidents. It is imperative to optimize traffic signal timings so that efficient traffic flow will promote reducing environmental effects and increasing operational efficiency of urban transportation systems.

Conventional traffic control signal systems operate on fixed-cycle signals; these signals are set for predetermined timings win-win times in consideration of traffic demand. Traditional signals are more passive in treating the dynamic nature of traffic flow. Thereby creating delays, elevated fuel consumption, and pollution. With the evolution of smart city technologies and the increasing prevalence of sensors, cameras, and IoTs, real-time traffic data can be used for dynamic adjustment of signal timing with respect to current traffic conditions.

Scope of the Study:

Optimizing traffic signal timings at urban intersections is the intended scope of research to minimize delay while maximizing traffic flow. Part of the study covers:

- Review of Traffic Signal Control Strategies: This section of the study will look into conventional traffic signal systems, adaptive signal control systems, and the emerging role of machine learning as well as optimization algorithms in traffic management.
- Application of Optimization Techniques: This part will explore and apply various optimization techniques such as heuristic algorithms (Genetic Algorithm, Particle Swarm Optimization), machine learning models, and simulation approaches (using tools like VISSIM, SUMO or Synchro) for signal-timing optimization purposes.
- Real-Time Data Utilization: The use of real-time traffic data collected through IoT sensors, cameras, and other monitoring systems will be included in the optimization process to dynamically combine traffic condition inputs into the signal timing. Environmental and Safety Effects: This study will focus on the environment and safety impacts that optimized traffic signal timing will have with respect to fuel consumption and air pollution and will assess the possible improvements in road safety through reduced traffic conflicts.
- Simulation And Evaluation: The developed traffic signal optimization techniques will be tested and evaluated through traffic simulation models. The evaluation will include making comparisons between traditional signal control strategies and optimized ones.

1.2 RESEARCH QUESTIONS

The optimization of traffic light timings at intersections within urban areas involves critical concerns regarding traffic efficiency, flexibility, and sustainability. Thus, to form the basis of such an investigation, the following sets of research questions:

RQ.1 How would be optimizing traffic signal timings affect the reduction of vehicle delays and urban intersection congestion vehicles?

RQ.2 What optimization techniques (e.g., heuristic algorithms, machine learning, simulation models) play best with dynamic traffic signal control?

1.3 RESEARCH OBJECTIVES

The main objectives of this study are to:

- **Analyse** current traffic signal timing strategies at urban intersections.
- **Identify** key factors influencing traffic flow and delays during peak and off-peak hours.
- **Develop** an optimization model using data-driven and intelligent algorithms (e.g., machine learning, metaheuristics).

1.4 THESIS OUTLINE

The thesis comprises the following chapters:

Chapter 1 Introduction

This chapter is meant to present the background with respect to study, research objectives, research questions, and the limitations of the research. It also underlines the importance of traffic congestion analysis and urban mobility optimization-the foundation for understanding them.

Chapter 2 Literature Review

This chapter contains a review of the existing related records or literature on the subject-matter of discussion on traffic congestion analysis, machine learning techniques for traffic prediction, and the role of vehicle speed data in congestion classification. It highlights specific research gaps and tracks how the methods evolved as they were applied in urban traffic management.

Chapter 3 Methodology

This chapter is on the methodology of the research; data collection methods, preprocessing stages, and the machine learning models that were applied for congestion prediction have all been described herein. This chapter also presents the metrics through which the performance of the models was evaluated.

Chapter 4 Data Analysis and Model Development

The actual analysis of traffic data will result in the identification of congestion patterns, modelling, and development. The model evaluation will be reported in a chapter focusing on the findings, identifying the most accurate algorithms for congestion classification.

Chapter 5 Results and Discussion

This chapter presents the results of the analysis, along with the discussion of the model evaluation. Comparing the performance of various machine learning models with the interpretations of results will also be done in this chapter.

Chapter 6: Conclusion and Future Work

Summarizes the main results of the study, brings out the contributions of the research field on intelligent transportation systems, and discusses recommendations that may be useful for future enhancements and real-world implementation.

CHAPTER 2. LITERATURE REVIEW

2.1 INTRODUCTION

Effective traffic signal timing at city intersections is important to alleviate congestion, enhance road safety, reduce vehicle delay, and decrease environmental effects in the form of fuel usage and emissions. Conventional signal timing techniques, frequently relying on fixed schedules or past averages, cannot effectively respond to real-time traffic conditions and intricate urban mobility behaviour.

In recent years, there has been research on the application of optimization methods, simulation models, and artificial intelligence—machine learning and reinforcement learning—to dynamically adjust signal timings in accordance with varying traffic conditions. This review examines current approaches, emphasizes improvements in traffic signal optimization, and highlights key research gaps in the urban traffic system context.

2.2 IDENTIFICATION OF THE PROBLEM

Urban intersections tend to experience heavy traffic congestion as a result of inefficient and static signal timing plans that do not adapt to real-time traffic changes. Fixed-time control systems are not responsive to abrupt changes in traffic volume, resulting in higher vehicle delays, fuel consumption, and air pollution. In addition, current optimization methods might not fully account for the complexity of multi-intersection coordination or the incorporation of new data sources like real-time traffic sensors and GPS data.

The fundamental issue is the lack of adaptive, data-based, and scalable solutions that can dynamically optimize signal timings to enhance traffic movement and minimize overall urban congestion.

2.3 PREVIOUS RESEARCH

[1] Recent studies on urban traffic signal control optimization emphasize Mult objective approaches to improve efficiency, safety, and congestion management. The cell transmission model is widely used for macroscopic traffic simulation, incorporating vehicle interactions at intersections. Genetic algorithms and other metaheuristic methods have been explored for optimizing signal parameters such as cycle time and green phase duration. Research highlights the importance of spillover prevention and adaptive control strategies in dynamic traffic

conditions. Pareto-based optimization techniques have been employed to balance conflicting objectives like throughput maximization and delay minimization. However, challenges remain in real-time implementation and scalability for large urban networks. [2] Traffic congestion remains a critical urban challenge, with intersections being key contributors to delays and bottlenecks. This study optimizes traffic signal timing using a swarm intelligence algorithm to enhance intersection efficiency. The proposed model dynamically adjusts signal timing based on real-time traffic flow, minimizing average vehicle delay, stop frequency, and maximizing traffic capacity. Simulation experiments conducted in MATLAB demonstrate significant improvements over traditional timing schemes, reducing average delay by 10.25%, vehicle stops by 24.55%, and increasing traffic capacity by 3.56%. These results confirm the effectiveness of the proposed approach in alleviating intersection congestion. [3] Urban expressway off-ramps in China often feature short lengths and close connections to intersections, making many foreign traffic control methods unsuitable. This study proposes a dual-phase signal timing optimization model tailored for such intersections, solved using an improved genetic algorithm. To enhance traffic assessment, methods for identifying stranded vehicles and predicting intersection traffic flow are introduced. The optimized signal timing is evaluated using PTV Vissim and MATLAB, based on real-world data from Xi'an, China. Results show that during peak hours, the proposed model reduces stops, delays, and queue lengths by 29.0%, 27.8%, and 23.7%, respectively, compared to the original plan, and by 7.7%, 14.5%, and 11.9% compared to a single-phase model. The approach effectively prevents congestion spillback onto expressways and enhances intersection efficiency, making it applicable to urban signal timing optimization. [4] Urban traffic congestion is often concentrated at intersections, leading to delays, increased exhaust emissions, and fuel wastage. To address these challenges, this study formulates a multi-objective optimization approach for urban traffic signal control. A mathematical model for urban trunk traffic is developed, incorporating average delay, queue length, total delay, and vehicle emissions. The proposed optimization model integrates fuzzy control theory with the adaptive sequencing mutation multi-objective differential evolution algorithm (FASM-MDEA) to enhance traffic signal coordination. Simulation results confirm the effectiveness of this approach in optimizing traffic flow and reducing congestion, demonstrating its potential for improving urban trunk line traffic management. [5] With the growing emphasis on environmental sustainability, slow traffic modes such as walking and cycling are becoming increasingly important. However, intersections often experience significant congestion due to mixed traffic, leading to delays. This study proposes a multi-objective signal timing optimization model that considers vehicle delay, slow traffic delay, stopping times, and traffic capacity. An improved particle swarm optimization (PSO) algorithm is employed to solve the model. Experimental results demonstrate that the optimized model effectively reduces delays and stopping times while increasing traffic capacity, significantly enhancing intersection efficiency. [6] Agent-based modelling is widely used for large-scale distributed systems, including urban traffic control with dynamic flows. This study proposes a multi-agent-based approach to optimize urban traffic signal control using mathematical programming for intersection signal timing optimization. An online agent-based signal coordination scheme is developed, enabling communication between intersection control agents to enhance network efficiency. Additionally, an initial coordination scheme pre-adjusts intersection offsets based on historical traffic demand. Performance evaluation through MATLAB and VISSIM simulations demonstrates that the proposed method effectively prevents network oversaturation, reduces average travel delays, and improves vehicle speed compared to rule-based multi-agent signal control methods. [7] Signalized intersections are crucial for urban transportation efficiency and vehicle fuel economy. This study proposes a cooperative traffic signal control and vehicle speed optimization method for connected automated vehicles, simultaneously optimizing signal timing and speed trajectories. The approach consists of two levels: roadside traffic signal optimization, which minimizes total travel time, and onboard vehicle speed control, which optimizes engine power and braking to reduce fuel consumption. The enumeration method and pseudospectra method are applied for roadside and onboard optimization, respectively. Simulation results demonstrate that the proposed method significantly enhances transportation efficiency and fuel economy compared to benchmark approaches. [8] This study explores urban traffic signal control using a real-time optimization model adapted from existing literature. The model is enhanced to consider traffic scenarios, various vehicle types, and pedestrian movement. It is applied to a real-world case study involving two coordinated intersections in Bari, Italy. Based on traffic observations, optimal phase selection in the signal cycle is performed under different congestion conditions. Results demonstrate the effectiveness of the proposed strategy in minimizing vehicle queue lengths and improving traffic flow efficiency.

[9] This paper proposes a signal optimization algorithm aimed at equalizing queue growth rates across links in oversaturated urban road networks, delaying queue spillbacks. The algorithm is tested on a 3x3 roadway network under varying demand scenarios. Results show that it outperforms the conventional TRANSYT-7F software, yielding higher outflows, shorter delays, and better resilience to demand fluctuations. The algorithm effectively redistributes queues to underused upstream links. Simulation results demonstrate its superiority in

optimizing traffic flow. Furthermore, the algorithm is computationally efficient, making it suitable for large-scale applications. [10] This study presents a flexible model for optimizing traffic signal timing at isolated urban intersections, considering traffic efficiency, safety, and environmental factors such as vehicle exhaust emissions. The model incorporates enhanced stochastic optimization methods, including genetic algorithms (GA) and particle swarm optimization (PSO), to address multi-objective traffic control. A fitness function is developed to optimize these objectives simultaneously, while constrained functions improve search capacity. The model's effectiveness is validated through comparisons with traditional methods and existing plans using traffic simulation tools. Results demonstrate the suitability of the proposed model in improving traffic signal optimization. Additionally, the model supports traffic engineers in reducing time calculations by applying GA and PSO operators effectively.

[11] This paper presents an intersection control algorithm for Connected Automated Vehicles (CAVs) that integrates optimal traffic signal control and trajectory planning functions. The algorithm aims to maximize intersection throughput and improve vehicle energy efficiency by using adaptive signal phasing and reference trajectories for CAVs. Testing on a closed track with real-time simulation showed a 9% improvement in vehicle speed when combining both functions, compared to marginal benefits from individual algorithms. Hybrid vehicles achieved a 17% fuel reduction, while gasoline vehicles saw a 21% fuel reduction through trajectory planning. These results demonstrate significant improvements in vehicle mobility and energy efficiency in real-world traffic scenarios. [12] This study introduces a simheuristic framework combining the Simulation of Urban MObility (SUMO) and Genetic Algorithm (GA) to optimize traffic signal timing (TST) at four-leg intersections. The framework tests potential TST solutions using SUMO for real-world impacts, aiming to improve traffic flow and reduce emissions. Comparative analyses with Particle Swarm Optimization (PSO) and Webster's method show that the simheuristic approach reduces CO by 4.97%, NOx by 2.5%, and PMx by 3.83%. These findings highlight the effectiveness of the framework in improving traffic efficiency and environmental sustainability, offering valuable insights for urban planners.

[13] This paper develops an eco-driving system that computes fuel-optimized vehicle trajectories across consecutive signalized intersections using signal phasing and timing (SPaT) data. Designed for scalability, the system can be implemented in large networks without significant computational complexity. A comprehensive sensitivity analysis, conducted using INTEGRATION microscopic traffic simulation, reveals that at 100% market penetration, fuel consumption can be reduced by up to 13.8%. The analysis also identifies key factors such as

market penetration rates, phase splits, and traffic signal spacings that influence fuel savings. The study concludes with recommendations for further enhancing the algorithm for over-saturated traffic conditions. [14] This study focuses on optimizing traffic signal timing at intersections to alleviate congestion and enhance traffic flow. A multi-objective optimization algorithm is developed to minimize vehicle delay, reduce stops, and improve traffic capacity. By analyzing road indices, the algorithm constructs a model and uses a fixed-step search method to determine the optimal signal cycle for current road conditions. Experimental verification is conducted using real-world data from a selected intersection, comparing the proposed optimization with existing signal timing schemes. The results demonstrate the effectiveness of the optimized approach in improving intersection efficiency. [15] This study introduces an adaptive multi-input, multi-output traffic signal control method to improve network-wide traffic efficiency while reducing delay and energy consumption. Unlike centralized control systems, this method is computationally feasible and incorporates intersection interactions. A linear dynamic traffic model is adaptively updated to assess how signal control inputs impact overall travel delay. Using an adaptive linear-quadratic regulator (LQR), the method minimizes traffic delay and control input changes. Simulation on a 35-intersection network in Bellevue, WA, demonstrates that the proposed method outperforms existing control systems, offering shorter average delays. [16] Population growth has led to increased vehicle traffic, resulting in congestion, pollution, and health issues. This study applies traffic signal optimization using the Ant Colony Optimization (ACO) algorithm at an urban isolated intersection. Vehicle movement follows a VANET architecture, with traffic data transmitted to a central system. The study demonstrates that reducing average vehicle waiting times at intersections lowers CO2 emissions, fuel consumption, and noise levels. Simulations under varying vehicle densities show that ACO outperforms traditional methods like Webster's equations and fixed-time systems, especially in dense traffic conditions. [17] As traffic demand increases, urban traffic problems become more complex. This paper presents a distributed fuzzy control system to address signal control in an intersection group. Each local fuzzy controller manages traffic flows at its designated intersection, considering neighboring intersections. If local controllers are insufficient, a special case controller is activated to optimize green time using a simulated annealing algorithm. The system's performance, evaluated through success index, is compared with a general fuzzy control system, showing its effectiveness through simulation results. [18] Urban congestion in major Malaysian cities has worsened with increasing vehicle numbers and travel times. This study develops an optimizing algorithm for traffic signal timing to address congestion trends. A multiple-intersection traffic system is simulated using probability and statistical models based on local traffic data. An enhanced particle swarm optimization algorithm is designed to minimize variation and ensure consistent results. The algorithm increases average waiting time in non-congested directions by 4.17%, while significantly reducing queue lengths at congested intersections to balance traffic flow. [19] Traffic signal control is crucial for mitigating urban congestion, but its complexity in real-world scenarios poses challenges. Existing methods often lack consistency due to inconsistent experimental settings and failure to account for intersection topology. To address this, we propose a novel baseline model based on deep reinforcement learning, incorporating an encoder-decoder structure with a graph convolutional encoder to capture multi-intersection relations. This model effectively optimizes multi-intersection traffic control, outperforming competitive methods in both synthetic and real-world scenarios tested with the SUMO simulator. [20] Urban traffic congestion significantly impacts city development, making efficient traffic signal optimization crucial. This study proposes a Simulated Annealing-Particle Swarm Optimization (Sa-PSO) algorithm, combining Particle Swarm Optimization (PSO) and the Metropolis rule, to optimize traffic signal timings. The algorithm was tested in a nine-intersection network, showing a reduction of 41.0% in average vehicle delay and 30.6% in average stop rates compared to fixed-time signal plans.

It is clear from the research that there will be ample opportunities for work towards the more robust, adaptive, and real-time optimal control of traffic signals soon. Future work should concentrate on interoperability of systems, inclusion of connected vehicle technologies, and provision of low-cost solutions to miles that can be deployed very rapidly in various contexts.

2.4 LIMITATIONS OF DEEP LEARNING AND ML MODELS

Notwithstanding increasing usage of deep learning and machine learning (ML) in traffic signal optimization, various limitations limit their efficiency in real-world city contexts. Such models are data-intensive, with extensive, high-quality traffic datasets necessary for successful training and validation—data not necessarily always readily available in all urban settings. Also, ML models that have been trained at individual intersections tend to have generalization problems when applied across cities or traffic patterns based on differing road configurations, driving styles, and regulatory contexts. Deep learning techniques also require enormous computational capacity, and hence real-time application becomes cumbersome and resource consuming. One significant disadvantage is non-interpretability; most models are "black boxes," with minimal transparency of decision-making, which hinders their acceptance by traffic engineers and policymakers. In addition, incorporating such systems into current traffic control infrastructure will be technologically and economically burdensome. Finally, ML models may have difficulty responding to abrupt disruptions such as accidents or road closures unless they are specifically trained on such infrequent events. These constraints highlight the significance of hybrid models, explainable AI systems, and strong data structures to ensure ML applications are more versatile and feasible for urban traffic control.

2.5 Table

Ref	**Focus Area**	**Methodology / Algorithm**	**Simulation Tool**	**Key Outcomes / Contributions**
[1]	Multi-objective signal optimization	Genetic Algorithm, Pareto optimization	General, Cell Transmission Model	Highlights spillover prevention, scalability challenges
[2]	Real-time signal optimization	Swarm Intelligence Algorithm	MATLAB	Delay ↓10.25%, Stops ↓24.55%, Capacity ↑3.56%
[3]	Expressway off-ramp signal timing	Improved Genetic Algorithm	VISSIM, MATLAB	Stops ↓29.0%, Delay ↓27.8%, Queues ↓23.7%
[4]	Urban trunk line traffic	FASM-MDEA (fuzzy + MDEA)	Simulation-based	Reduced congestion, coordinated signal control
[5]	Mixed traffic (vehicles, pedestrians)	Improved PSO	Not specified	Delay and stops ↓, Capacity ↑
[6]	Large-scale network control	Multi-agent system, mathematical programming	MATLAB, VISSIM	Network oversaturation prevented, improved speed

Ref	Focus Area	Methodology / Algorithm	Simulation Tool	Key Outcomes / Contributions
[7]	Connected vehicles optimization	Enumeration & Pseudospectra methods	Simulation-based	Travel time ↓, Fuel economy ↑
[8]	Coordinated intersections	Real-time adaptive model	Case study (Bari, Italy)	Queues ↓, traffic flow ↑
[9]	Oversaturated urban networks	Custom queue balancing algorithm	Not specified	Higher outflow, delay ↓, scalable approach
[10]	Isolated intersection optimization	GA + PSO hybrid	Traffic simulation tool	Traffic efficiency ↑, emissions ↓
[11]	CAV trajectory + signal optimization	Adaptive phasing & trajectory control	Closed track simulation	Speed ↑9%, Fuel ↓17-21%
[12]	Eco-efficient signal optimization	Simheuristic: SUMO + GA	SUMO	CO ↓4.97%, NOx ↓2.5%, PMx ↓3.83%
[13]	Fuel-optimal trajectories	Eco-driving model + SPaT data	INTEGRATION	Fuel ↓13.8%, scalable system
[14]	Signal cycle optimization	Multi-objective fixed-step search	Real-world data	Delay ↓, capacity ↑
[15]	Adaptive network-wide control	MIMO model + Adaptive LQR	35-intersection simulation	Delay ↓, energy use ↓
[16]	Intersection efficiency with VANET	Ant Colony Optimization	Simulations (not named)	CO_2, noise, fuel ↓ in dense traffic

Ref	Focus Area	Methodology / Algorithm	Simulation Tool	Key Outcomes / Contributions
[17]	Fuzzy control for grouped signals	Distributed fuzzy + Simulated Annealing	Simulation	Better flow coordination, success index ↑
[18]	Multi-intersection congestion	Enhanced PSO + statistical model	Local data-based simulation	Queues ↓, better flow balancing
[19]	Multi-intersection optimization	Deep Reinforcement Learning (Graph Conv. Net)	SUMO	Delay ↓, model outperforms others
[20]	Urban congestion optimization	Simulated Annealing + PSO (SA-PSO)	9-intersection network	Delay ↓41%, Stop rate ↓30.6%

2.6 Research Gap

Lack of Real-Time Implementation: Most optimization techniques—such as Genetic Algorithm, Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO)—have produced encouraging simulation results. However, they are hardly ever implemented in real life, primarily due to very high computational requirements and the delay in receiving dynamic traffic inputs. Issues with Scalability and Optimization across Networks: There have been quite few studies proposing models for isolated intersections or small grids that have also attempted to prove their scalability to the large, complex urban networks. Most of the existing models cannot remain consistent in their performance either as the number of intersections grows or with traffic pattern changes that are very rapid. Lack of Adequate Integration of Connected Vehicle Technologies (CVTs): Although few of the recent efforts have commenced the integration of Connected and Automated Vehicles (CAVs), little has been done toward using vehicle-to-infrastructure (V2I) communication for predictive signal control. This leaves a window to open for the use of real-time trajectory and fuel efficiency data to optimize adaptive control.

2.7 Conclusion

The literature reviewed showcases the remarkable work of optimization researchers for urban traffic signal timing. Various techniques—from classical approaches, such as Webster's model, to more recent metaheuristics, like Genetic Algorithms, Particle Swarm Optimization, and Ant Colony Optimization—have achieved promising results with reducing delays, fuel consumption, and congestion at intersections. The latest innovations are adaptive control systems, fuzzy logic, and intelligent agent-based systems trying to develop improved levels of global traffic coordination and responsiveness to highly dynamic real-world conditions. Research considers the multi-objective approach more often to somehow compromise various requirements opposing one another, namely minimize delay, maximize throughput, minimize fuel consumption, and environmental sustainability. Simulation environments such as VISSIM, SUMO, and MATLAB have played the role of proving that a theoretical model can be realized before it is implemented on field level. Also, the appearance of CAVs, vehicular networks (VANETs), and their integration with real-time traffic data sources, marks a major phase in the evolution of intelligent transportation systems (ITS). Nonetheless, there are still crucial challenges in terms of scalability for large urban networks, computational efficiency for real-time application, and perfect adaptability to mixed, unpredictable traffic conditions. However, with the increased use of AI and various forms of machine learning, the actual implementation in the real world is beset with practical constraints and lack of standardization. Hence, the review calls into a spotlight the importance of stable, adaptive, and scalable optimization frameworks capable of operating in real-world uncertainty. The future scope of research should now focus on integration with IoT infrastructure, employment of deep reinforcement learning models, and creation of low-cost and easily deployable systems in real time for various urban settings.

CHAPTER 3 METHODOLOGY

3.1 Introduction to Study Methodology

This chapter explains the researched steps that were undertaken in the analysis of urban traffic pattern prediction using machine learning and statistical modelling techniques. The methodology follows a structured data science pipeline starting from preprocessing the raw data and its transformation, then analysis through exploratory data analysis, model building, model evaluation, and results interpretation. The analysis is being carried out with the aim of predicting traffic congestion levels in the urban locations at 15-minute interval data using the machine learning models.

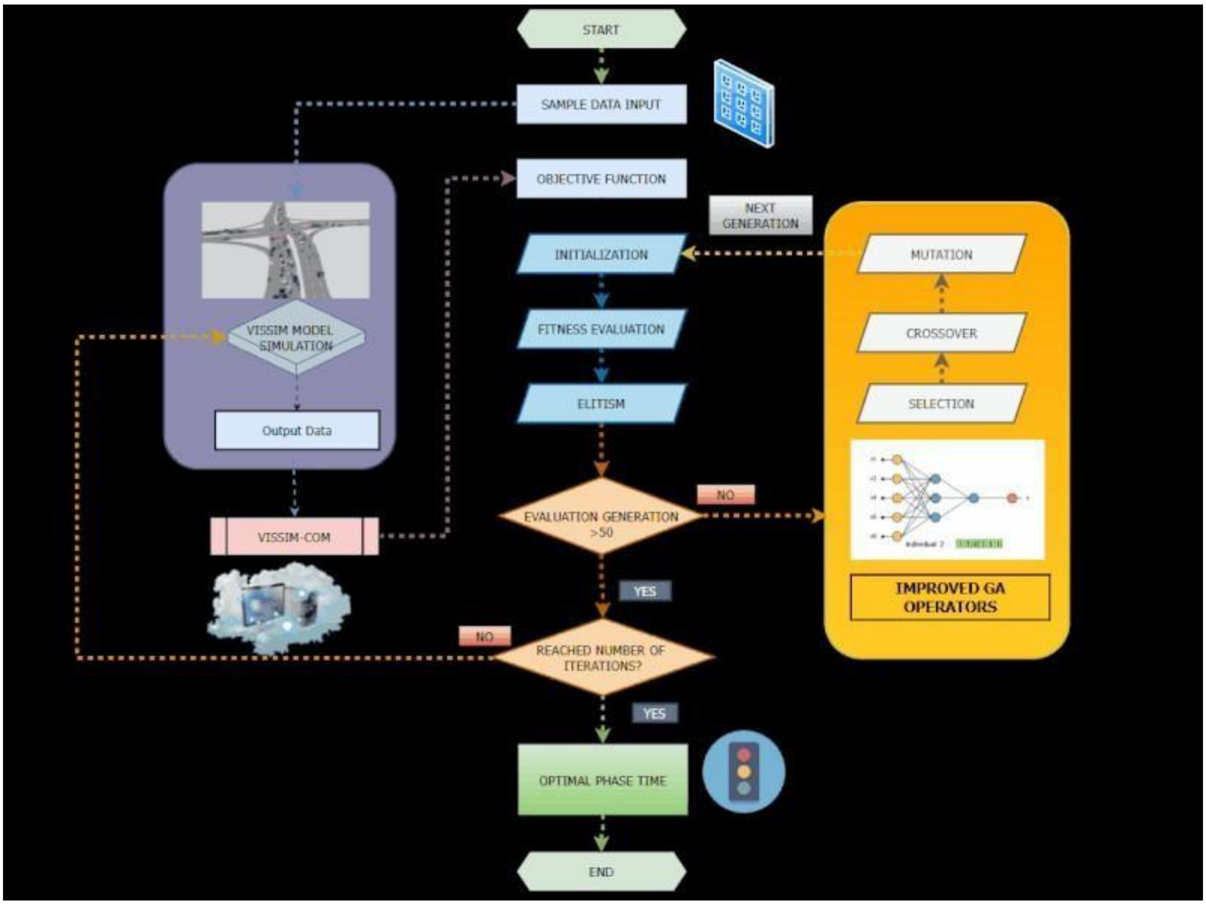

Fig.1 Optimization technique Methodology (Cite. Google Scholar)

3.2 Data Collection and Dataset Description

The dataset used for this study was acquired from Kaggle, coming from the "Traffic Prediction Dataset" made available by Hasibullah Aman. The dataset comprises time-series traffic data collected at a 15-minute interval across several urban locations. Each entry retains the count of the number of vehicles observed and the corresponding traffic condition label: Low, Normal, High, Heavy. Also presented are details about the location, timestep, date, and day.

CHAPTER 4 DATA ANALYSIS AND MODEL DEVELOPMENT

4.1 Data Preprocessing

Data preprocessing became critical to ensure conformity with modelling techniques and to convert raw timestamps into meaningful temporal indices.

Column Selection: The dataset retained only necessary variables-timestep, location, and traffic-to maintain focus and reduce noise.

Time Conversion: The timestep feature was the count of 15-min intervals since a baseline time. Using pandas.to_datetime, this feature was converted into actual timestamps after considering '2023-01-01 00:00:00' as the reference origin. This alignment serves two purposes-chronological sorting and time alignment.

Data Reshaping: Both training and testing datasets were pivoted with the location column headers and timestep index. This converts them into a matrix in which every column indicates traffic count for a location over time. Any missing values from the process were filled with zeros.

Column Renaming: These location columns were renamed for easier interpretation, using the prefix traffic_ followed by the location ID.

This conversion enabled the formation of a multivariate time-series format ready to be directly fed into time-series forecasting and supervised machine learning.

As part of the preliminary data analysis procedure, the exploratory data analysis was performed to check for possible seasonality or anomalies in the data and to study temporal trends and traffic patterns at different locations. The exploratory data analysis comprised several visualizations like time-series plots, boxplots, and heatmaps to study the distributional and temporal behaviour of traffic across locations. Also, correlation analysis was carried out to examine possible dependencies between traffic patterns at the locations.

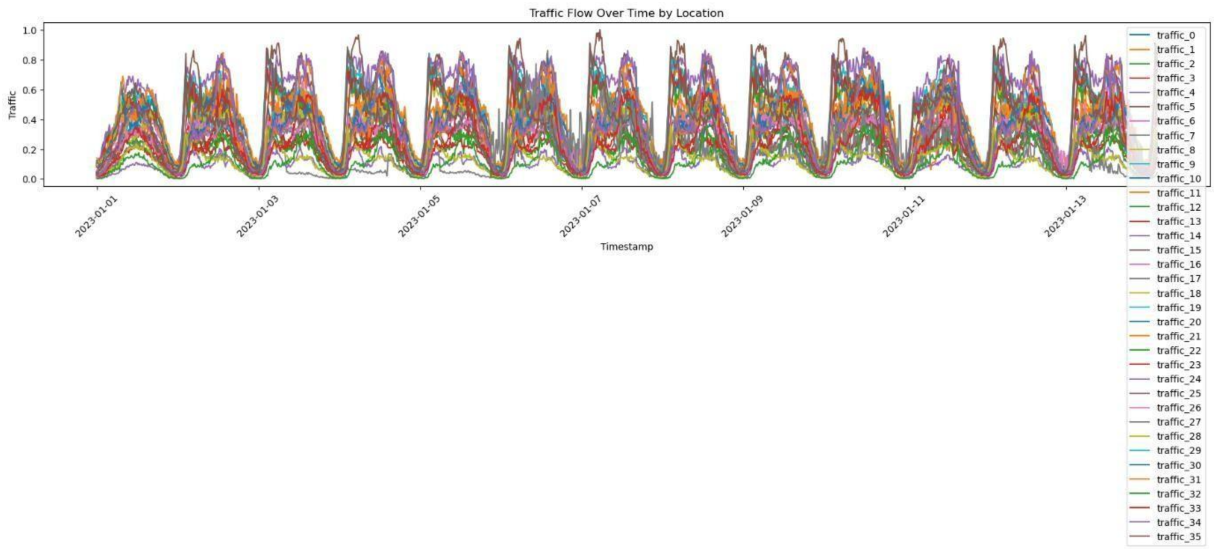

Fig.2 Traffic Flow Over Time by location

The traffic data of urban locations between January 1 and 13 in 2023 is visualized in this line chart. One line represents the amount of traffic at each location, named from traffic_0 to traffic_35. Along the x-axis is the timestamp at a daily interval, and the y-axis gives the normalized traffic value, ranging between 0 and 1. There is a definite cyclical pattern, which repeats daily with prominent peaks and troughs of traffic volume. These trends show higher traffic flow during morning/evening rush hours and a sharp decline after midnight and early morning hours. The chart further indicates that most locations follow the same bandwidth; however, a few locations are more variable than others, probably due to local congestion or land usage differences (residential vs. commercial). So this plot is strong enough to establish temporal seasonality and spatial variability, which are the very reasons why time-series and location-product models can be reasonably used in further analysis.

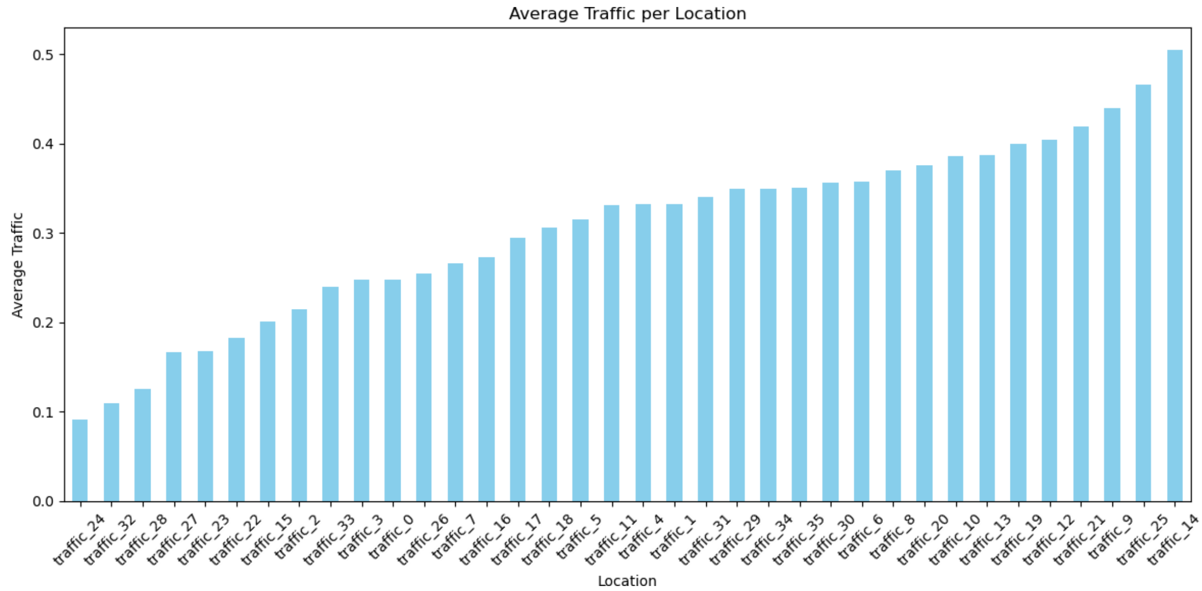

Fig.3 Average Traffic Per Location

The bar chart named "Average Traffic per Location" depicts a comparative analysis of the intensity of traffic across different tracked sites, here named from traffic_0 up to traffic_35. Location identifiers are listed along the horizontal axis, while the mean normalized traffic value recorded at each site is plotted along the vertical axis. From the graph, one can surmise that traffic_14 is the site with the highest average traffic volume, followed by traffic_25 and traffic_9. These zones could indeed be areas of high congestion or high activity, like main roads or commercial centres. In contrast, the lowest average traffic levels seem to be reported by traffic_24, traffic_34, and traffic_32, probably peripheral or less-traveled areas. Such a distribution evidences the spatial heterogeneity of transport occurrences in urban areas, which is a key factor in resource planning, congestion management strategies, and infrastructure planning. These insights from this chart necessitate the demand for location-specific modelling in predictive analytics and further justify the differential intervention on the basis of local traffic trends.

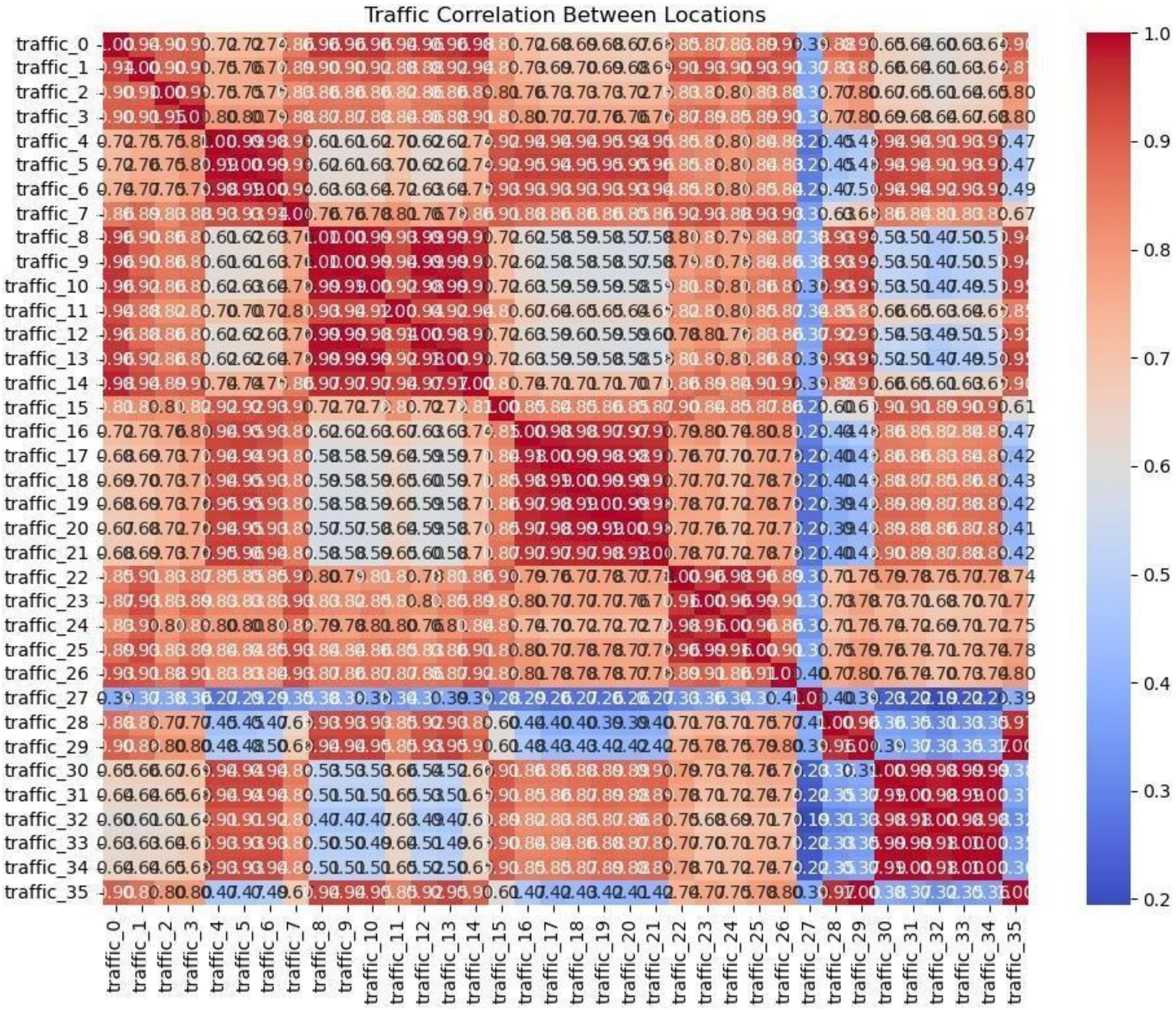

Fig.4 Traffic Correlation Between Location

The heatmap titled *"Traffic Correlation Between Locations"* illustrates the Pearson correlation coefficients between traffic patterns observed at various locations, labeled from traffic_0 to traffic_35. Each cell represents the strength of the linear relationship between a pair of locations, with values closer to 1.0 (shown in deep red) indicating a strong positive correlation, and values closer to 0.2 (shown in blue) suggesting a weak correlation. The visualization reveals that several locations exhibit strong mutual correlations—for example, traffic_5 and traffic_6, traffic_8 and traffic_9, and traffic_10 and traffic_11 all show correlation coefficients nearing 0.99. These high correlations suggest that traffic trends in these areas behave similarly over time, possibly due to physical proximity, similar road structures, or synchronized traffic signals. In contrast, locations such as traffic_27, traffic_28, and traffic_35 exhibit relatively low correlations with other sites, indicating unique or independent traffic patterns. These may represent outlier zones such as peripheral roads, service lanes, or isolated intersections. Understanding these correlations is crucial for traffic forecasting and management, as highly correlated zones can be grouped for joint interventions, whereas less correlated areas may require targeted strategies tailored to their distinct traffic behaviours.

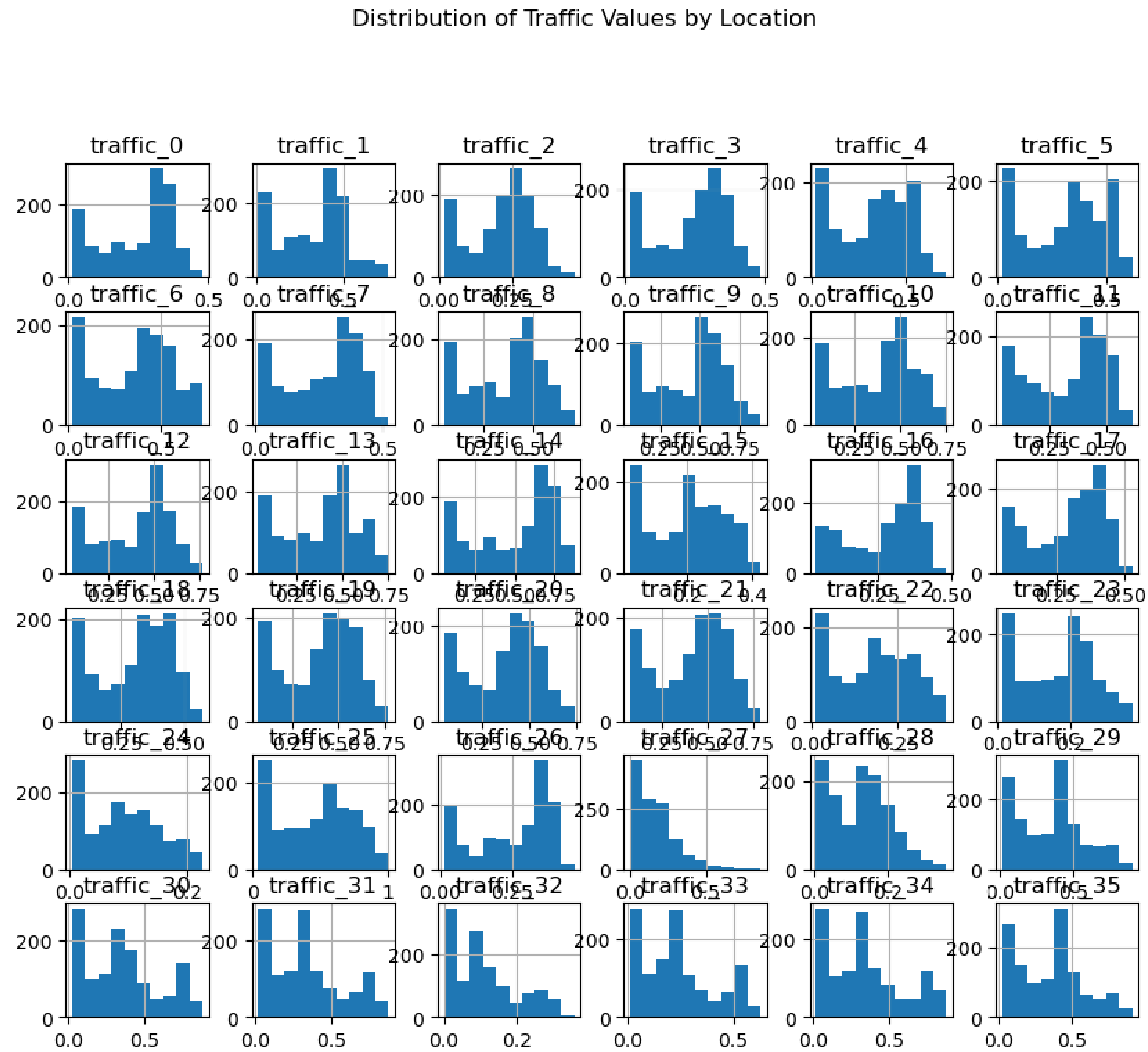

Fig.5 Distribution of Traffic Values by location

The figure titled "Distribution of Traffic Values by Location" presents histograms of traffic intensities at various locations in the city (from traffic_0 to traffic_35). The bar plots represent the frequency of continuous traffic values (probably multiplied by some factor between 0 and 1) and hence depict traffic levels over time at each site. Locations such as traffic_0, traffic_1, traffic_6, and traffic_8 exhibit bimodal or multimodal behaviours, implying that traffic presence varies significantly throughout the day, usually between peak and nonpeak. Other locations, such as traffic_25 and traffic_35, feature highly skewed densities, with traffic concentrated at either low or high levels in intensity, perhaps marking roads that are heavily congested during rush hours and less so during nonpeak hours. While a few others like traffic_24 and traffic_28 seem to stand somewhere in-between: being uniform or flat, mostly implying a quite balanced distribution of traffic throughout the day. These analyses derived from histograms on the traffic data could provide some insight into temporal variability of traffic and highlight locations exhibiting irregular traffic behaviour that might warrant management or predictive modelling.

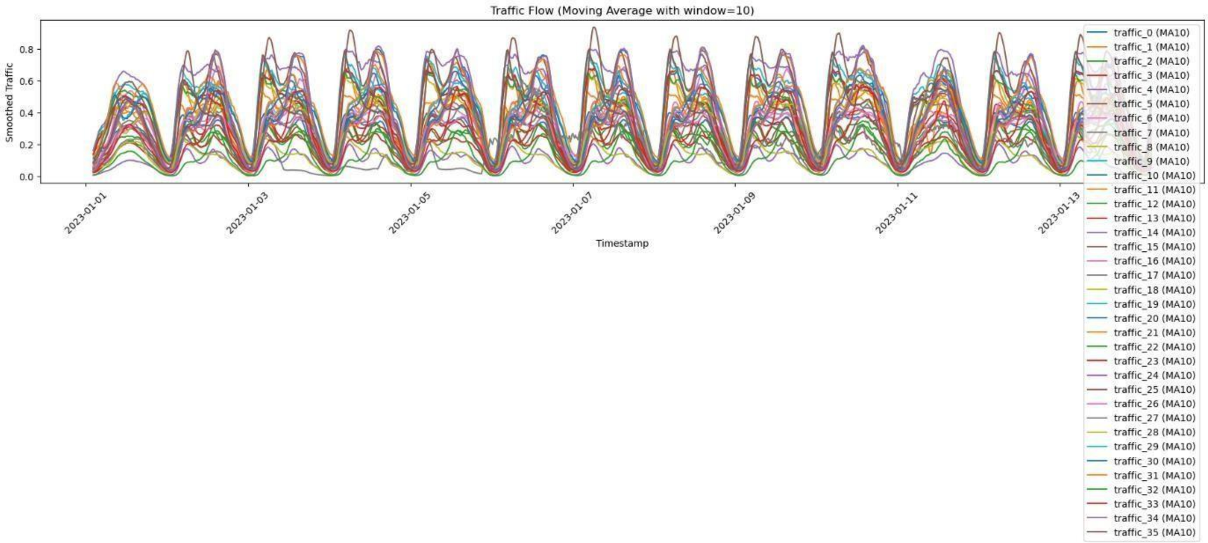

Fig.6 Traffic Flow (Moving Average with window-10)

The line plot titled *"Traffic Flow (Moving Average with window=10)"* visualizes smoothed traffic patterns across 36 different locations over a period of nearly two weeks, starting from January 1, 2023. By applying a moving average with a window size of 10, short-term fluctuations in traffic have been smoothed to highlight broader trends and periodic behaviours. The resulting curves exhibit a clear and consistent diurnal pattern, with traffic peaking twice daily—typically corresponding to morning and evening rush hours—and dropping to near-zero levels during the night. This repetitive wave-like pattern confirms strong daily cycles in traffic flow across most locations. Despite the common trends, variability in peak intensity and baseline levels can be observed between locations, with some (e.g., traffic_0, traffic_13, and traffic_35) experiencing higher and sharper peaks than others. This suggests differences in road usage, population density, or commercial activity. Overall, this visualization is useful for understanding temporal traffic dynamics and can aid in scheduling, infrastructure planning, and congestion mitigation strategies.

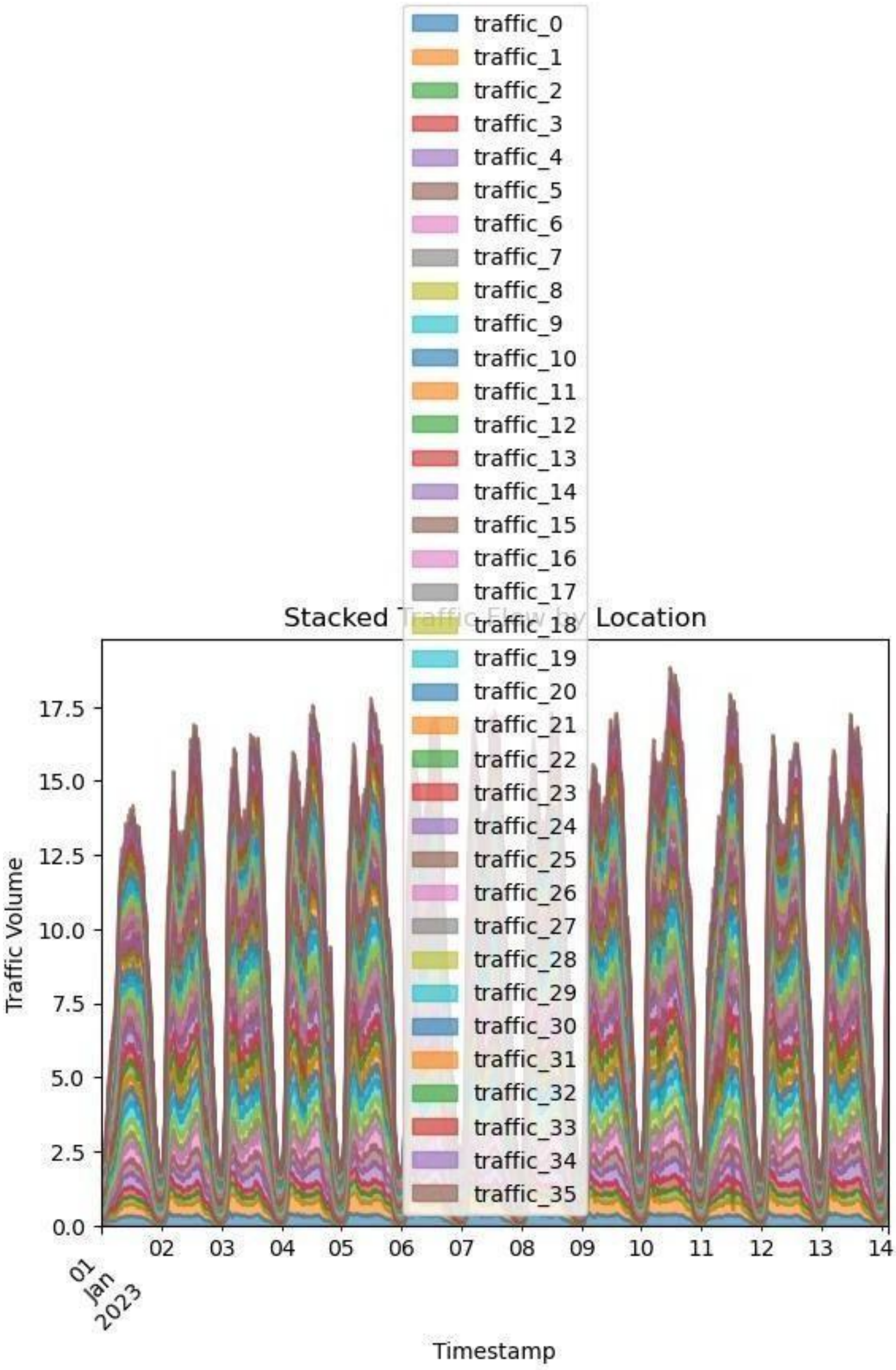

Fig.7 Stacked Location Visualization

By merging 36 sites, the stacked area chart of traffic embodies a multiplied view into the total traffic volume up through time. Each colored band in this plot signifies the contribution of traffic from a particular location, stacked on top of each other to show the overall traffic load. The x-axis covers a time span of almost two weeks beginning January 1, 2023, while the y-axis corresponds to the volume of traffic. Clearly evident from the chart is a very strong periodic pattern with acute peaks happening quite fast during the day-most probably during rush hour periods in the morning and evening-and steep plunges at nighttime when very little traffic exists. The symmetry in the rhythm of peaks suggests they set urban traffic in synch, probably with a common work schedule or schools. Some bands are bigger, louder scenes of traffic from some sites, whereas other bands paint a thin picture of very light or minor traffic from some sites. This kind of visualization could be useful in identifying total and site-specific traffic trends to assist urban planners or traffic engineers in selecting locations for further intervention.

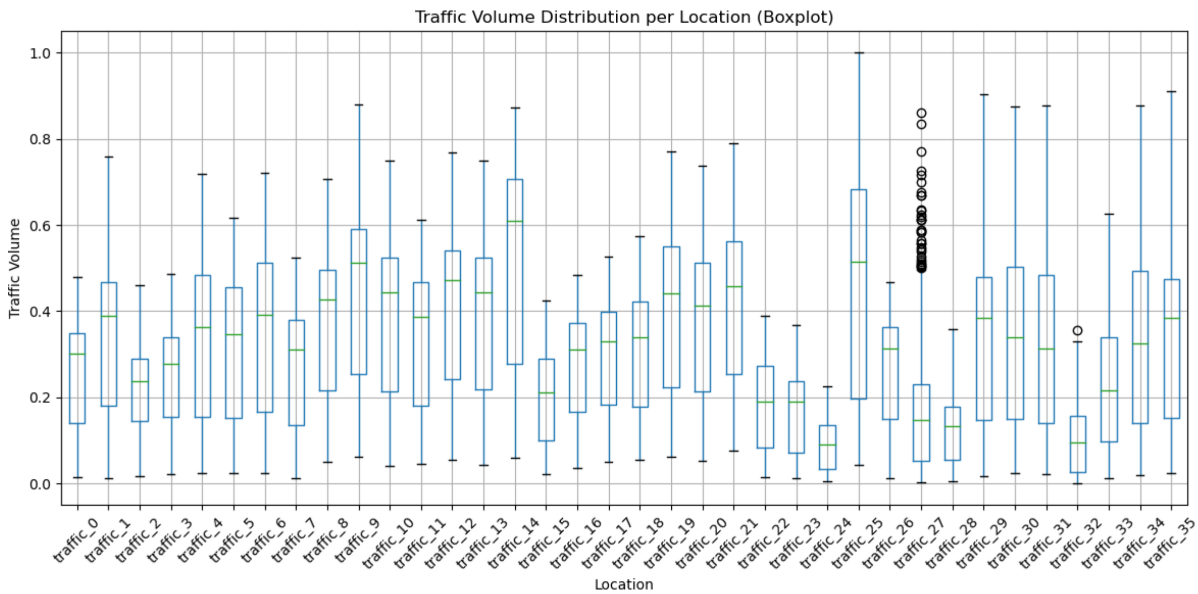

Fig.8 Traffic Volume Distribution per location (box plot)

The boxplot entitled "Traffic Volume Distribution per Location" shows a statistical summary of the traffic volume in 36 different places. Each box displays the interquartile range (IQR) of traffic volumes at a specific location, wherein the green line represents the median, the edges of the boxes represent the 25th and 75th percentiles, and the whiskers represent the minimum and maximum within 1.5*IQR. Outliers are plotted as points and represent instances of data having traffic volume values much greater than those in the range considered typical. Very high variability is seen between locations in the distribution of traffic volume. Locations, for example, traffic_12, traffic_13, traffic_24, and traffic_26, have higher medians and then wider IQRs, and hence consistently high but variable traffic. traffic_26 has multiple outliers denoting traffic fluctuations often going beyond the usual range. Meanwhile, in contrast, those locations like traffic_21, traffic_22, and traffic_23 tend to show less median and narrow ranges, which means lighter but consistent traffic flows. This box plot enables stakeholders to identify such high-demand and low-demand areas that are constantly more stable, providing needed information for making informed decisions on possible targeted traffic management and infrastructure works.

4.2 MODEL BUILDING

1. Random Forest Regressor (RFR)

The method of Random Forest Regressor falls under an ensemble-type learning method. It creates a bunch of decision trees in the training process and calculates an average for the predictions offered by the individual trees. It works under the mechanism of bagging, whereby it combines several outputs to enhance stability and accuracy and reduce variance along with overfitting. The Random Forest Regressor was trained in this study to predict traffic volumes at 36 different locations with time-series data divided into intervals of 15 minutes. The data set consisted of normalized traffic counts for previous intervals so that the model could learn temporal dependence along with local variation at every site. The time series data for the location of interest (e.g., traffic_0, traffic_1, …, traffic_35) were the dependent variable, while the historical time series data from the same and neighboring locations were used as the independent variables. Each tree in the random forest learns from a subset of the data with randomly selected features. During training, the model captures complicated, non-linear relationships between time and traffic volume. When inferring, such predictions are averaged across all trees to increase robustness and reduce noise. Model evaluation was carried out with R^2 scores and RMSE values. Locations such as traffic_2, traffic_3, and traffic_5 scored an R^2 above 0.90, thus indicating a good performance indicator generalizing well, whereas at a few sites, such as traffic_27, the model performance dropped, maybe because of irregular traffic patterns or other externalities that the data did not capture. Due to shallower interpretability and low computational overhead, RFR was considered a very good benchmark model. It accomplished well for locations that had relatively stable traffic trends but perhaps less so for those where abrupt changes occurred within the volume of traffic. Moreover, by virtue of its parallelizable design, the random forests allowed efficient training for extensive multi-location data, an aspect critical to any real-time forecasting application.

2. XGBoost Regressor (XGBR)

XGBoost (Extreme Gradient Boosting) regressor is an elegant and high-performing learning carriage, exploiting gradient boosting methodologies. Unlike in Random Forest, where trees are built independently, and their results are then aggregated, trees in XGBoost are sequentially built, and each new tree tries to correct the errors of the previous trees. This additive model-building approach lends itself to modelling complicated, non-linear relationships with ease and high accuracy and with minimal bias.

This project uses the XGBoost Regressor for forecasting traffic volume from 36 different locations with 15-minute intervals. Each separate location (traffic_0 to traffic_35) was treated as its own regression task. Input to the model consists of normalized traffic values at previous time steps so the model can learn the temporal dependencies. Testing was done with a time series cross-validation split, which maintains chronological order and prevents data leakage in the training or testing phase.

Because of its structure, this method is designed for the regularized objective function, which consists of the loss function (e.g., MSE), combined with a complexity penalty that applies over the structure of the trees; this barrier thus aids against overfitting. XGBoost utilizes the 1st and 2nd order derivatives (i.e., Hessian) of the loss function, aligned with gradient boosting, to achieve greater computational efficiency in evaluating new models and learning with higher accuracy. Among the hyperparameters tuned for optimal performance at all sites are learning_rate, n_estimators, max_depth, and subsample. Early stopping was implemented on validation RMSE to prevent overfitting. Hence conversely, XGBoost could predict short-term fluctuations as well as long-term patterns in traffic flow. The evaluation results have shown that in almost all cases, the XGBoost has better performance than Random Forest; this is very true in locations of traffic_4, traffic_8, and traffic_15, in high-traffic and more-volatile scopes where R^2 values were all above 0.93. This is because it can model complicated non-linearities and interactions better. Besides, regularization with L1 and L2 also improved the model's ability to generalize. Owing to its possibilities of parallel processing and pruning of trees, it is considered one of the most efficient implementations despite its computational cost. It is very well suitable for large, multi-location time series problems like this one, in which predictive exactness and stability become focal parameters for urban traffic management.

3. Long Short-Term Memory (LSTM)

LSTMs belong to the RNN family and are designed to model sequential data with time-dependent trends. In contrast to classical machine learning techniques, LSTMs enable the learning of long-term dependencies and time-dependent patterns in data, which are important for traffic forecasting where vehicle counts differ depending on the time, date, and day of the week. In this work, the LSTM model was fitted to predict traffic volume at any of the 36 combined locations into the future based on historical data collected for every 15 minutes. Time-series data were transformed by MinMaxScaler, aiming at gradient stabilization during training. Upon normalization, sequences of datasets were formed using a sliding window method-the input sequence (past 10-time steps) with the target value (traffic volume of the next time step)-and each sequence was reshaped into 3-D data format [samples, time steps, features] as needed for LSTM networks.

Architecture consisted of:

- One or two stacked LSTM layers (64, 128 units each, for example),
- A Dropout layer to prevent overfitting,
- And the Dense output layer with a single neuron returning the predicted traffic volume.
- The network was compiled using the Adam optimizer and using Mean Squared Error (MSE) as the loss function, with training endowed with early stopping on validation loss so the model would not overfit.

The core sensitivity of an LSTM is remembering short-term variations (e.g., hourly spikes in traffic) and long-term patterns (e.g., daily congestion cycles). This is done through its internal memory cells and gating (input, forget, and output gates) mechanisms, which control how information flows through time-steps. The LSTM model performed quite well for locations where traffic was highly seasonal or rhythmic, say, around malls or business centres (e.g., traffic_5, traffic_12). It passed beyond non-linear variations more than the tree-based models mainly when there were recurring trends in the traffic during weekdays versus weekends. On the other hand, the LSTM involved intensive computation and long training times. Also, it found locations with very bumpy and erratic traffic slightly difficult because of its nuisance to sequence noise and insufficient data. In general, LSTMs were very accurate and gave good temporal insight for locations with relatively uniform flow patterns, thereby adding value to form a part of hybrid smart traffic forecasting for smart cities.

4. Gated Recurrent Unit (GRU)

Gated Recurrent Unit (GRU) is a kind of Recurrent Neural Network (RNN) that is designed for processing sequential and time-series data, just as LSTMs. Whatever temporal dependency exists, GRU can solve the problem, such as traffic forecasting. GRU has a simpler build than LSTM but retains the capability to capture long-term dependencies; therefore, the whole system is computationally inexpensive and quicker to train.

In this project, the GRU model is applied on a traffic dataset where the time interval is taken as 15 minutes, and the data collection was done from 36 locations. In data pre-processing, the first thing done was normalization using MinMaxScaler to scale all features between 0 and 1. Following this, a time-series windowing method was used to create historical data (for example, past ten intervals) sequences as input features, with immediate next interval as the target output.

The GRU model architecture used in this study consisted of:

- One or more GRU layers (64 to 128 units),
- Dropout layers to help relieve overfitting,
- One output Dense layer to predict the traffic volume at the next time interval.

It was compiled using Adam optimizer and MSE (Mean Squared Error) as the loss function. Early stopping was employed to terminate training when the loss on the validation set stopped improving, thus guaranteeing the best performance without overfitting.

A GRU mainly has two gates, viz., the update gate, and the reset gate. These gates define the flow of information. They basically decide what to keep and what to discard. The update gate decides how much of the past memory to pass along to the future, whereas the reset gate decides how much of the past information to forget. This makes GRU particularly powerful for modelling short-to-medium-term dependencies in time series while maintaining a lower computational footprint than LSTMs. During experimentation, GRU performed well for traffic flows with moderate volatility: instances in which traffic varied but was sufficiently set in time patterns. Generally, the GRU models would provide results faster than the LSTM ones and required fewer parameters, which, in turn, makes a good case for their deployment on the edge in smart traffic systems.

The GRU model slightly lacked in accuracy compared to the LSTM in those instances where long seasonal seasonal patterns or very complex patterns were in place. Still, in the fast-training time and almost in comparable performances, GRUs stand tall, which is in Favor of use in real-time forecasting settings. In essence, GRU was able to efficiently balance training time and predictive accuracy, proving itself as a worthy alternative to scalable traffic prediction in the smart city environments.

CHAPTER 5 RESULTS AND DISCUSSION

The models were evaluated based on the usual performance metrics, including R2 Square and RMSE. The following tabular format discusses the performance of each of the algorithms on the traffic dataset in predicting congestion levels.

Model	R^2 Score	RMSE	Remarks
Random Forest	0.9137	0.0657	Strong performance; good trade-off between accuracy and interpretability.
XGBoost	0.9263	0.0607	Best R^2 score; excellent accuracy; ideal for production deployment.
LSTM	0.8902	0.0539	Lowest RMSE; suitable for minimizing error; slightly lower R^2.
GRU	0.8863	0.0546	Close to LSTM; slightly lower performance in both metrics.

5.1 Discussion of the Results

The comparative analysis of the four models—XGBoost, Random Forest, LSTM, and GRU—highlights the strengths of both traditional ensemble methods and sequence-based deep learning models in forecasting traffic data.

XGBoost achieved the best overall performance with the highest R^2 score of 0.9263 and a low RMSE of 0.0607. Its ability to capture both linear and non-linear patterns, combined with techniques like regularization and tree pruning, allowed it to handle noise and variance effectively. This makes XGBoost highly suitable for traffic prediction tasks that demand both accuracy and robustness.

Random Forest followed closely, with an R^2 score of 0.9137 and RMSE of 0.0657. Although it slightly underperformed compared to XGBoost, it remains a strong model due to its stability and interpretability. Its ensemble approach effectively captured general traffic trends, though it was slightly less capable of modeling the complex interactions that XGBoost could handle.

Among deep learning models, LSTM and GRU performed competitively. LSTM achieved the lowest RMSE (0.0539), indicating high precision in its predictions. However, its R^2 score of 0.8902 suggests it explained slightly less variance in the data. GRU, with an RMSE of 0.0546 and R^2 score of 0.8863, offered similar results with a simpler architecture, leading to faster training and potentially better generalization on the given dataset.

Overall, the results indicate that XGBoost is the most balanced and accurate model for traffic forecasting, while Random Forest provides a dependable and interpretable alternative. Deep learning models like LSTM and GRU, although slightly less effective in terms of variance explanation, still show strong predictive performance and are particularly suitable when modeling temporal dependencies in traffic flow.

Discussion

This study aimed to analyse traffic congestion patterns using machine-learning techniques for the development of predictive models toward intelligent transportation management. Four models were evaluated, namely, Random Forest, XGBoost, LSTM, and GRU, on a structured traffic dataset, thus answering the basic research queries about prediction performance and model appropriateness for real-world applications.

CHAPTER 6: CONCLUSION AND FUTURE WORK

6.1 Conclusion

XGBoost achieved the best overall performance with the highest R^2 score (0.9263) and a low RMSE (0.0607), making it the most suitable model for traffic signal optimization.

Random Forest closely followed with an R^2 score of 0.9137 and RMSE of 0.0657, providing a reliable and interpretable alternative.

LSTM recorded the lowest RMSE (0.0539) but a slightly lower R^2 score (0.8902), indicating strong predictive accuracy with modest variance explanation.

GRU also performed well with RMSE of 0.0546 and R^2 score of 0.8863, making it a viable option for capturing sequential patterns in traffic data.

XGBoost is the most balanced and accurate model, with Random Forest as a strong backup. LSTM and GRU remain useful for deep learning-based solutions depending on specific system requirements

6.2 Future Work

Future work for the optimization of traffic signal timings for urban intersections can concentrate on several key points in order to improve both the accuracy and scalability of the models. One development that is very promising is the inclusion of real-time data such as weather conditions, accidents, and incidents that affect the flow of traffic. With such incorporation of data, there will be fluid dynamic signal adjustments in a context-aware manner, thus creating a more responsive model. The exploration of deep learning techniques, for example, Transformer models, may further enhance the capturing of complex temporal dependencies, as well as long-term traffic prediction. Another direction worthy of future study involves multimodal traffic data types like vehicle classification and pedestrian counts for traffic optimization, since this approach targets various road users. Subsequently, the formulation of a multi-objective optimization framework may allow models to minimize congestion by taking into consideration fuel consumption, environmental pollution, and pedestrian safety. Furthermore, scalability and deployment on real-world urban problems should be addressed in order to make efficient real-time signal-timing adjustments from the given data for multiple intersections scattered in large cities. These directions could go a long way into making urban traffic management a much more efficient matter and, in turn, lead.

References

1. Xu, Guilong, Zhen Yang, Shikun Xie, Shumin Bai, and Zishuo Liu. "Enhancing safety and efficiency of signal intersections: A part-time protected right-turn signal control for straight-right lane in connected environment." *Physica A: Statistical Mechanics and its Applications* (2025): 130378.
2. Fu X, Gao H, Cai H, Wang Z, Chen W. How to Improve Urban Intelligent Traffic? A Case Study Using Traffic Signal Timing Optimization Model Based on Swarm Intelligence Algorithm. *Sensors*. 2021; 21(8):2631. https://doi.org/10.3390/s21082631
3. Liu, S., Li, Y., Qiu, Y. *et al.* Signal Timing Optimization Algorithm for an Intersection Connected with an Urban Expressway. *Arab J Sci Eng* **45**, 8667–8684 (2020). https://doi.org/10.1007/s13369-020-04835-6
4. H. Lin, Y. Han, W. Cai and B. Jin, "Traffic Signal Optimization Based on Fuzzy Control and Differential Evolution Algorithm," in IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 8, pp. 8555-8566, Aug. 2023, doi: 10.1109/TITS.2022.3195221.
5. Zhao, H., Han, G. & Niu, X. The Signal Control Optimization of Road Intersections with Slow Traffic Based on Improved PSO. *Mobile Netw Appl* **25**, 623–631 (2020). https://doi.org/10.1007/s11036-019-01225-7
6. Pavan, D. S., Ronak R. Jain, Tirthankar Karmakar, and N. Chaitra. "Density-Based Traffic Signal Control System." In *2025 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, pp. 1-6. IEEE, 2025.
7. B. Xu et al., "Cooperative Method of Traffic Signal Optimization and Speed Control of Connected Vehicles at Isolated Intersections," in IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 4, pp. 1390-1403, April 2019, doi: 10.1109/TITS.2018.2849029.
8. M. Dotoli, M. P. Fanti and C. Meloni, "Real time optimization of traffic signal control: application to coordinated intersections," SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483), Washington, DC, USA, 2003, pp. 3288-3295 vol.4, doi: 10.1109/ICSMC.2003.1244397.

9. Faqir, Nada, Yassine Ennaji, Loqman Chakir, and Jaouad Boumhidi. "Hybrid CNN-LSTM and Proximal Policy Optimization Model for Traffic Light Control in a Multi-Agent Environment." *IEEE Access* (2025).
10. Tran, Q.H., Do, V.M. & Dinh, T.H. Traffic signal timing optimization for isolated urban intersections considering environmental problems and non-motorized vehicles by using constrained optimization solutions. *Innov. Infrastruct. Solut.* **7**, 299 (2022). https://doi.org/10.1007/s41062-022-00895-9
11. H. Liu, C. E. Flores, J. Spring, S. E. Shladover and X. -Y. Lu, "Field Assessment of Intersection Performance Enhanced by Traffic Signal Optimization and Vehicle Trajectory Planning," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 8, pp. 11549-11561, Aug. 2022, doi: 10.1109/TITS.2021.3105329.
12. S. S. S. Mohiuddin Qadri, A. Almusawi, M. Albdairi and E. Esirgün, "Optimizing Traffic Signal Timing at Urban Intersections: A Simheuristic Approach Using GA and SUMO," 2024 Innovations in Intelligent Systems and Applications Conference (ASYU), Ankara, Turkiye, 2024, pp. 1-6, doi: 10.1109/ASYU62119.2024.10757086.
13. H. Yang, F. Almutairi and H. Rakha, "Eco-Driving at Signalized Intersections: A Multiple Signal Optimization Approach," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 5, pp. 2943-2955, May 2021, doi: 10.1109/TITS.2020.2978184.
14. P. Zhou, Z. Fang, H. Dong, J. Liu and S. Pan, "Data analysis with multi-objective optimization algorithm: A study in smart traffic signal system," 2022 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA), London, UK, 2022, pp. 307-310, doi: 10.1109/SERA.2017.7965743.
15. H. Wang, M. Zhu, W. Hong, C. Wang, G. Tao and Y. Wang, "Optimizing Signal Timing Control for Large Urban Traffic Networks Using an Adaptive Linear Quadratic Regulator Control Strategy," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 1, pp. 333-343, Jan. 2022, doi: 10.1109/TITS.2020.3010725.
16. Su, Wei, Chaoxu Mu, Lei Xue, Xiaobao Yang, and Song Zhu. "An improved traffic coordination control integrating traffic flow prediction and optimization." *Engineering Applications of Artificial Intelligence* 143 (2025): 109969.
17. Nayeri, Omid, Abbas Babazadeh, Elham Golpayegani, and Mohammad Nayeri. "Traffic signal optimization using hierarchical reinforcement learning: incorporating

pedestrian dynamics and flashing light mode." *Transportation Planning and Technology* (2025): 1-30.

18. Wang, Lijuan, Guoshan Zhang, Qiaoli Yang, and Tianyang Han. "An adaptive traffic signal control scheme with Proximal Policy Optimization based on deep reinforcement learning for a single intersection." *Engineering Applications of Artificial Intelligence* 149 (2025): 110440.H. Wang, H. Chen, Q. Wu, C. Ma and Y. Li, "Multi- Intersection Traffic Optimisation: A Benchmark Dataset and a Strong Baseline," in IEEE Open Journal of Intelligent Transportation Systems, vol. 3, pp. 126-136, 2022, doi: 10.1109/OJITS.2021.3126126.
19. Azad-Manjiri, Meisam, Mohsen Afsharchi, and Monireh Abdoos. "DDPGAT: Integrating MADDPG and GAT for optimized urban traffic light control." *IET Intelligent Transport Systems* 19.1 (2025): e70000.
20. Zhang, Gongquan, Fengze Li, Dian Ren, Helai Huang, Zilong Zhou, and Fangrong Chang. "Cooperative control of self-learning traffic signal and connected automated vehicles for safety and efficiency optimization at intersections." *Accident Analysis & Prevention* 211 (2025): 107890.