Author/s:
Chan, LS;Nassir, N;Zhang, X;Yazdani, M;Sarvi, M

Title:
Preemptive crash risk reduction through a real-time cost-based safety prediction model (RECOSAM) for traffic signal control

Date:
2025-12-01

Citation:
Chan, L. S., Nassir, N., Zhang, X., Yazdani, M. & Sarvi, M. (2025). Preemptive crash risk reduction through a real-time cost-based safety prediction model (RECOSAM) for traffic signal control. Computers and Electrical Engineering, 128, https://doi.org/10.1016/j.compeleceng.2025.110639.

Persistent Link:
https://hdl.handle.net/11343/357972

Computers and Electrical Engineering 128 (2025) 110639

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Preemptive crash risk reduction through a real-time cost-based safety prediction model (RECOSAM) for traffic signal control

Lok Sang Chan *, Neema Nassir *, Xiaocai Zhang, Mobin Yazdani, Majid Sarvi

Faculty of Engineering & Information Technology, The University of Melbourne, VIC, 3010, Australia

ARTICLE INFO

Keywords:
Preemptive safety measures
Cost-based safety evaluation
Traffic optimisation
Multi-objective traffic signal
Machine learning

ABSTRACT

This paper proposes a novel real-time cost-based safety prediction model (RECOSAM) and incorporating it in intersection traffic signal control optimisation, complementing the recent advances in deep reinforcement learning (RL)-based adaptive traffic signal control (ATSC). The primary contribution is the development of RECOSAM, a model designed to predict traffic safety risks one step ahead of time, for various signal phase configurations at intersections. The proposed model offers a dynamic safety evaluation strategy, estimating near-future safety metrics for seamless integration into machine learning-based ATSC systems. Extensive experiments validate the model's effectiveness, demonstrating its potential for adaptive adjustments to mitigate impending safety risks. Perhaps more importantly from an operational policy perspective, the proposed model is capable of finding an optimal and justifiable trade-off between the efficiency of traffic flow and its safety in real-time.

A case study showcases the integration of RECOSAM into deep RL for green time optimisation. Results suggest that extended dedicated right turn phases may reduce safety risks, while overly protected phases could lead to inefficiencies in green time allocation and increased congestion. The model's adaptability across different scenarios is further illustrated, showing its capability to evaluate critical trade-offs between safety and efficiency especially for vehicles trying to make a right turn by finding gaps through traffic coming form the opposing direction (in left-hand-side driving countries—same applies for left turns in right-hand-side driving countries).

1. Introduction

In 2016, traffic crashes led to 1.35 million fatalities and over 50 million injuries globally [1], costing countries about 3% of their GDP [2]. Despite efforts to improve road safety, progress in reducing road traffic mortality has stalled. This stagnation urged a comprehensive approach towards identifying and evaluating potential crash risks within the transport system.

Urban intersections are vital components of the transport system and require special attention to ensure safety and efficiency. The challenges arising from the conflicting movements inside the intersection area. Recent deep reinforcement learning (Deep RL) based adaptive traffic signal control (ATSC) models have shown promise in optimising intersection performance, reducing delays, travel times, and queues [3–6]. In addition to optimising efficiency, multi-objective approaches have integrated safety and sustainability [7,8]. Urban transport is undergoing a significant transformation driven by population growth and the adoption of

* Corresponding authors.
E-mail addresses: loksangc@student.unimelb.edu.au (L.S. Chan), neema.nassir@unimelb.edu.au (N. Nassir).

https://doi.org/10.1016/j.compeleceng.2025.110639
Received 18 October 2024; Received in revised form 7 July 2025; Accepted 18 August 2025
Available online 31 August 2025

advanced transport modes. This urged careful operational management to mitigate potential injuries and fatalities. Although safety and efficient are distinct objectives, they are interlinked and should be cooperatively implemented for optimal traffic operations.

Traditional approaches to traffic safety rely on implementing mitigation measures at the construction stage in geometric design and later in traffic law enforcement, such as speed control and red/amber light violations. However, more in-depth operational measures, especially at intersection traffic signals are attracting attention recently [9,10].

Previous studies have modelled collision risk using real time traffic data [11]. However, several challenges remain, including lengthy data collection processes [12], under-reporting of crashes [13], and limited insight into the underlying causes of traffic incidents [14]. Assessing traffic conflicts as a surrogate for collisions has emerged as a promising alternative, offering more frequent observations and the ability to capture near misses that bridge the gap between crashes and non-crashes [12,15]. With the advancement of connected vehicle technology and high resolution data, real time risk assessment has become increasingly feasible [16].

Nonetheless, obtaining real time safety insights alone is not sufficient for their integration into machine learning based traffic signal control systems. Existing safety evaluation frameworks often rely on step by step simulations and external software, making them computationally intensive and unsuitable for real time applications. Additionally, many reinforcement learning based traffic signal control models either neglect safety entirely or consider it in a qualitative or threshold-based manner, without providing a unified and quantifiable metric that can be optimised alongside efficiency objectives. There remains a critical need for a safety evaluation model that is both computationally efficient and capable of providing meaningful and comparable safety feedback in real time.

In this study, we propose a real-time cost-based safety prediction model (RECOSAM) considering variable demands and different phase configuration for traffic operations in real settings, which has the potential for implementation in real-time decision-making and optimisation. Our key contributions are listed below:

1. We developed RECOSAM to estimate near-future safety performance for traffic signal control optimisation. By converting safety risks into cost, considering both economic and human costs associated with crashes, RECOSAM provides a comprehensive framework for evaluating trade-offs between safety and efficiency at intersections.
2. For safety risk evaluation, we consider both the probability of a crash event and its severity in traffic conflicts. RECOSAM is the first model to explicitly takes into account these two key risk components, providing a more robust and justifiable assessment of safety risk in various traffic scenarios.

The paper is structured as follows. Section 2 reviews real-time safety evaluation models and multi-objective RL-based signal control. Section 3 provides details on the data collection process and model training for RECOSAM. Section 4 presents the experimental results and a case study, with Section 4.1 exploring RECOSAM's variability, Section 4.2 examining the effects of diverse traffic operations across varying traffic environments through a preliminary safety optimisation algorithm, and Section 4.3 demonstrating its integration into a deep reinforcement learning framework. Finally, Section 5 concludes the paper, and Section 6 outlines the limitations of the current work and suggests directions for future research.

2. Literature review

The following literature review addresses four key areas of progress in traffic safety and adaptive signal control systems. It begins with advancements in real-time safety risk evaluation at signalised intersections, focusing on both collision-based and conflict-based models. The shift from historical crash data to traffic conflicts as a data source is also discussed, highlighting the role of connected vehicle technologies in improving data resolution. The review then explores the integration of reinforcement learning (RL) with deep learning and the evolution towards multi-objective optimisation frameworks that incorporate objectives such as safety and sustainability.

2.1. Real-time collision-based evaluation model

Extensive research has focused on modelling real-time collision risk. Studies by Wang et al. [17] and Khattak et al. [11] estimate safety performance functions (SPFs) for urban signalised intersections, accounting for crash types and severity. These studies utilise crash data to assess risk and predict crash occurrences based on traffic variables. Additionally, the advancements in real-time data collection have enabled prediction of vehicle occupant injury risks [18]. Hu et al. [19] mapped historical crash occurrences using connected vehicle data and deep learning models, such as multi-layer perceptron (MLP) and convolution neural network (CNN), showcasing their use in identifying high-risk intersections.

While collision-based models are critical for safety assessment, crashes are infrequent, leading to lengthy data collection and limited insights into accident causes [12–14].

2.2. Traffic conflicts as a surrogate of crashes

Alternatively, using traffic conflicts as a crash surrogate has gained attention due to its higher frequency, offering more detailed insights. Zheng et al. [20] reviewed this correlation, finding strong connections in multiple studies. Charly and Mathew [21] developed a lane-based model showing temporal correlations with historical crashes. Peesapati et al. [22] highlighted Post-Encroachment Time (PET) as an effective metric with low thresholds, while Zheng and Sayed [23] created a Bayesian model using indicators like modified Time to Collision (MTTC) and PET for crash estimation. Subsequently, Fu and Sayed [24] used Bayesian modelling to estimate real-time crash risk from MTTC-based conflicts, showing strong alignment with historical crash data and highlighting conflicts' predictive value.

Analysing conflicts reduces reliance on crash data, providing richer insights into hazardous movements since conflicts occur more frequently. Guo et al. [25] introduced Bayesian Tobit models for real-time rear-end conflict rate SPFs, addressing risk ambiguity from variable cycle lengths.

Hu et al. [26] explored the link between traffic variables and conflicts using a lane-based real-time safety evaluation model, illustrating the value of conflicts for vehicle-based safety assessment.

Lastly, Hussain et al. [27] proposed a unified framework combining extreme value theory (EVT) and autoregressive models to forecast crash risk at intersections. The EVT model captured dynamic traffic conditions, while the moving average model predicted crash risk up to 30 min ahead, demonstrating the framework's predictive capability.

2.3. Deep reinforcement learning in traffic signal control

The use of RL for traffic signal control started with the SARSA algorithm for single intersections [28]. In the 2000s, researchers advanced to network-level coordination using Q-learning [29–33].

Recently, integrating deep learning with RL has enabled more complex traffic control systems. The introduction of the Deep Q-Network (DQN) by DeepMind [34] marked significant progress. Genders and Razavi [35] expanded DQN's capabilities by using dense raw traffic data (DTSE), while Garg et al. [36] and Sajad Mousavi et al. [37] improved training stability with the Policy Gradient (PG) algorithm.

Further innovations include Oroojlooy et al. [38], which used attention mechanisms to handle complex intersection geometries. Devailly et al. [3] showed deep RL's scalability to large networks, and Li et al. [4] improved multi-intersection coordination via knowledge sharing. Wang et al. [5] developed a cooperative multi-agent framework for communication among intersection groups, demonstrating that shared global rewards improve convergence over independent RL approaches.

2.4. Multi-objective deep RL-based adaptive traffic signal control

In conjunction with the evolution of RL-based adaptive traffic signal control, considerable attention has also been directed towards multi-objective optimisation in recent years. Traditional intersection optimisation methods focused exclusively on traffic efficiency in the control objectives [10]. However, the dynamic shifts in urban transport, coupled with increasing popularity of micro-mobility options and the rise of new paradigms for inclusive multimodal transport, have caused concerns for road user safety [15,19,25,26,39] and operational sustainability [40–42]. Consequently, a sequence of studies has emerged with the aim of incorporating safety or sustainability considerations into traffic signal control [7–10].

Khamis et al. [7] developed a multi-agent RL traffic signal control model that integrates multiple objectives, adapting to changing environments. They assigned weights to objectives, enabling representation through a single rewards function. The model addressed congestion and considered weather-induced impacts on speed parameters affecting traffic safety.

Stevanovic et al. [8] devised a multi-objective optimisation model employing an evolutionary algorithm, encompassing considerations of traffic mobility, surrogate safety, and environmental factors. Results demonstrated the model's superior performance against original signal time setting from research field observations. Li and Sun [10], presented a multi-objective optimisation methodology for signal setting and lane allocation at intersections. Their approach embraced objectives including transport efficiency, road safety, and energy economy. The collective findings of these studies provide valuable insights into the integration of multi-objective considerations when overseeing traffic at intersections.

Lastly, Du et al. [9] introduced the SafeLight framework, which integrates a safety model into an RL model by Liang et al. [43]. The model assesses the safety of actions based on current traffic conditions, with a focus on left turn motion in left-hand driving scenarios, a common source of crossing collisions. Various approaches were experimented with to integrate the safety model into state-of-the-art RL models. Results demonstrate that integrating safety into deep RL traffic signal control enhances performance across various RL models, outperforming alternative collision minimisation methods.

The proposed RECOSAM model presented in this paper advances the state of the art in multiobjective traffic signal control. None of the existing models reviewed quantify and incorporate safety risk in terms of a cost value that can be directly balanced against time savings. Moreover, existing studies do not explicitly evaluate both key dimensions of risk, which are the probability of a crash and its potential severity. The proposed approach introduces a monetised risk metric that combines these dimensions into a single, interpretable measure. This enables consistent comparison with efficiency-based objectives and supports more informed decision making. The cost-based representation of safety is readily understandable by both road users and policy makers, offering a practical and justifiable tool for traffic management.

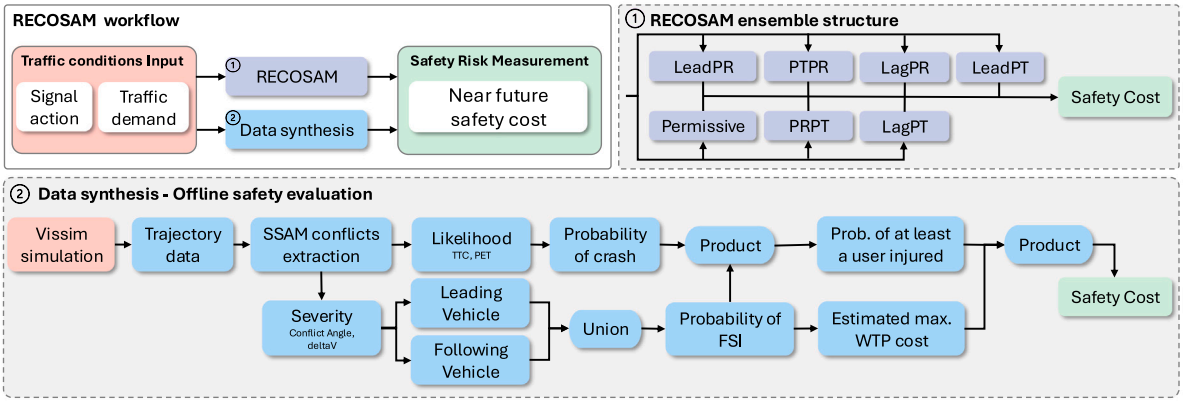

Fig. 1. RECOSAM workflow.

Table 1
Definition of preset phase configuration.

Phase configuration	Name	Abbr.	Sequence
$P_s = 1$	Permissive	Permissive	3-A-R
$P_s = 2$	Leading protected right	LeadPR	2-3-A-R
$P_s = 3$	Lagging protected right	LagPR	3-A-2-A-R
$P_s = 4$	Protected right protected through	PRPT	2-A-1-A-R
$P_s = 5$	Protected through protected right	PTPR	1-A-2-A-R
$P_s = 6$	Leading protected through	LeadPT	1-3-A-R
$P_s = 7$	Lagging protected through	LagPT	3-A-1-A-R

Note: In the sequence, the number represents the elementary phase id (E), while 'A' and 'R' denote the transition amber and all-red phases, respectively.

Despite recent progress, important gaps remain in the integration of safety within real time adaptive traffic signal control. Existing models based on historical crash data or surrogate safety measures often require intensive computation or external tools, limiting their applicability in real time environments. At the same time, most reinforcement learning based traffic control systems focus primarily on efficiency and rarely incorporate safety as a quantifiable and comparable objective. Where safety is included, it is often treated qualitatively or as a threshold condition, without providing a unified metric for optimisation. These limitations motivate the development of RECOSAM as a predictive, monetised safety model that supports real time safety-aware optimisation of traffic signal policies within a reinforcement learning framework.

3. Methodology

This section begins by outlining the data synthesis process for the training dataset of RECOSAM. The subsections start with a set of phase configurations along with simulation parameters designed with specific attention to the conflicting movement between right turns (left-hand-side driving) and opposing through movements. Following this, the offline safety evaluation pipeline, illustrated in Fig. 1 path (2), is presented. It comprises conflict identification using Surrogate Safety Measures (SSM), followed by cost quantification. Although the offline safety evaluation framework offers an easily understandable and justifiable measure of safety risk, it is unsuitable for real-time decision-making in ML-based ATSC due to its computational intensity and reliance on external software. In contrast, RECOSAM, as depicted in Fig. 1 path (1), is a simulation-based machine learning model that provides a predictive shortcut suitable for real-time ATSC decisions. The details of RECOSAM's model architecture are outlined in Section 3.2. It is noted that its integration into an adaptive control setting is discussed in the case study presented in Section 4.

3.1. Data synthesis

Continued research in traffic microsimulation show promising improvements in the accuracy and applicability of conflict-based safety evaluation [44–46]. In this study, a microsimulation using the mature software Vissim by the PTV Group was employed to generate trajectory data for training of RECOSAM. The simulation model was developed and calibrated based on an isolated intersection within the Australian Integrated Multimodal EcoSystems (AIMES) testbed in Carlton, Melbourne, Australia [6].

The signal control system incorporates a conceptual adaptive stage-based traffic signal that alternates between the north–south (NS) and east–west (EW) stages. Each stage (s) is composed of one or two of the three elementary phases: protected through, protected right, and permissive phases, as illustrated in Fig. 2. Given the different sequences of these elementary phases (E_k), where k denotes the positional index, there are a total of seven possible phase programs (P), as detailed in Table 1.

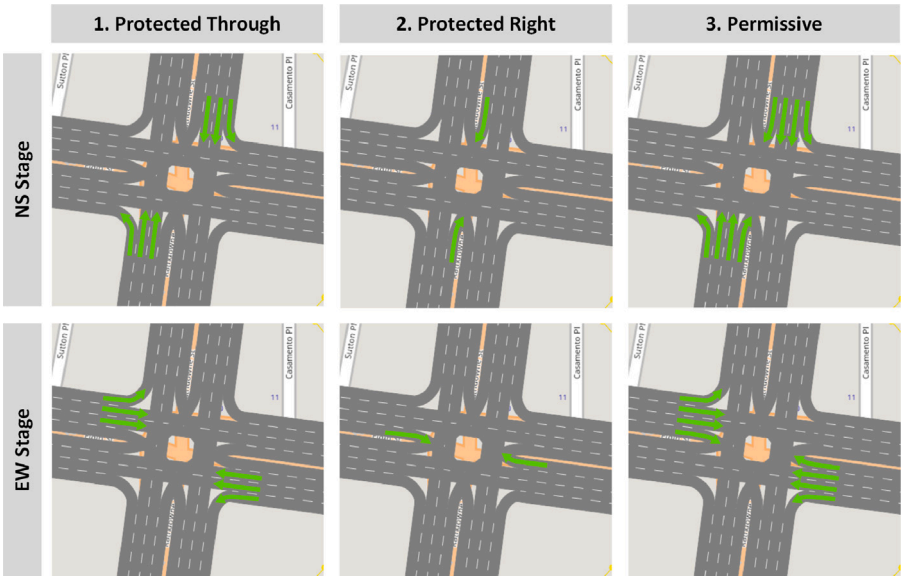

Fig. 2. Elementary phases (E_k).

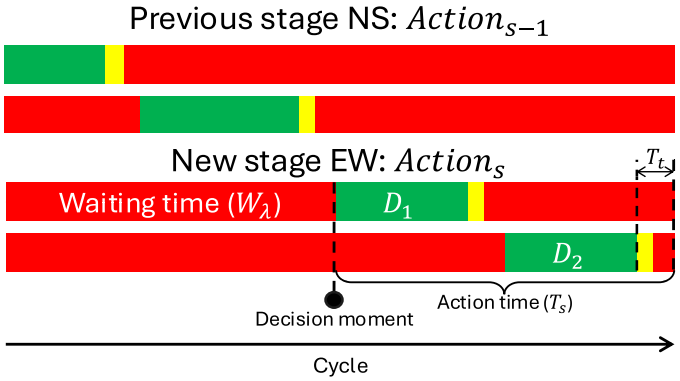

Fig. 3. Illustration of a stage-based two-phases traffic signal plan.

The signal control action, therefore, involves selecting one of these seven phase programs and determining the duration (D_k) for the corresponding elementary phases. Thus, the action for stage s can be represented as follows,

$$\text{Action}_s = \begin{bmatrix} P & D_1 & D_2 \end{bmatrix}_s . \tag{1}$$

The total effective time for the action (T_s), which encompasses the green duration of each elementary phase as well as the transition time (amber and all-red time), is expressed by

$$T_s = D_1 + T_t + D_2 + T_t, \tag{2}$$

where T_t represents the transition time following each elementary phase. Specifically, for the permissive phase program ($P = 1$), which contains only a single elementary phase, the action for stage s simplifies to

$$\text{Action}_s = \begin{bmatrix} 1 & D_1 & 0 \end{bmatrix}_s \tag{3}$$

and the corresponding action time becomes

$$T_s(P = 1) = D_1 + T_t. \tag{4}$$

The traffic demand parameters include traffic volume (O_λ), waiting time (W_λ) due to the previous stage ($s-1$), and the proportions of right ($\theta_{\lambda,1}$) and left ($\theta_{\lambda,2}$) turns, i.e.,

$$\text{Demand}_{\lambda,s-1} = \begin{bmatrix} O_\lambda & \theta_{\lambda,1} & \theta_{\lambda,2} & W_\lambda \end{bmatrix}_{s-1} , \tag{5}$$

Table 2
Definition of RECOSAM input variables.

Variable	Definition
Traffic signal parameters	
s	Active stage index, affects direction (iterative): if g is odd, NS (north-south), else EW (east–west)
P	Phase configuration id variable for each stage, taking values from 1 to 7 as presented in Table 1
k	Elementary phases position, k = 1 first phase; k = 2 s phase
E_k	Elementary phase id of position k, selectable from 1 to 3 as presented in Fig. 2
D_k	Elementary phase duration of position k, if P = 1, [5, 55] s, else [5, 25] s
Action_s	Action of stage $s = \begin{bmatrix} P & D_1 & D_2 \end{bmatrix}_s$
T_s	Action time of stage $s = D_1 + T_t + D_2 + T_t$
T_t	Transition time including amber and all-red time
Simulation traffic input parameters	
λ	Approach λ = 1, 2, 3, 4 corresponds to southbound, westbound, northbound and eastbound respectively
O_λ	Average vehicle volume per quarter hour of approach λ, [0, 500]/q h
$\theta_{\lambda,i}$	Proportion of turning vehicle i, i = 1 Right turn; i = 2 Left Turn, [0,100]%
W_λ	Waiting time due to previous stage, [10, 60] s
$\text{Demand}_{\lambda,s-1}$	Demand of approach λ from stage $s-1 = \begin{bmatrix} O_\lambda & \theta_{\lambda,1} & \theta_{\lambda,2} & W_\lambda \end{bmatrix}_{s-1}$

where λ represents the indices of each approach. In total, 16 parameters describe the traffic demands for all approaches. Fig. 3 illustrate an example of a signal cycle plan which consist of the NS and EW stages and the corresponding elementary phases. Waiting time (W_λ) is defined as the red time before the upcoming stage (s, EW in this example) since the activation of the previous stage ($s-1$, NS in this example). The north-south road is considered the major road, while the east–west road is designated as the minor road. To synthesis training data for RECOSAM, the traffic demands and signal control action are assigned randomly for each scenario. The effective range of each parameters is summarised in Table 2.

The Surrogate Safety Model (SSM) is employed to extract conflict events from simulated trajectories. This method has gained considerable attention in recent years, with numerous studies utilising microsimulation in conjunction with the Surrogate Safety Assessment Model (SSAM) software package [47–50]. The SSAM, published by the Federal Highway Administration (FHWA) [51] of the United States, is a prominent tool used for assessing SSMs from microsimulated trajectory data. In this approach, as illustrated in Fig. 1 path (2), multiple simulation instances are executed for each traffic scenario. The SSAM is then applied to analyse the trajectory information from microsimulations to identify and evaluate potential conflicts between vehicles. It estimates conflict indicators by employing surrogate safety measures, including Time to Collision (TTC) and Post-Encroachment Time (PET), to identify potential conflicts. TTC measures the time required for two vehicles to collide if there are no changes in speed or path [52], while PET is defined as the time difference between one vehicle leaving the potential collision area and another arriving [53].

The extraction of conflict data involves setting thresholds for indicators like TTC and PET, which vary across studies. Suggested thresholds range from 1.5 to 5 s for TTC and 1 to 4 s for PET [54,55]. To ensure a comprehensive inclusion of conflicts, threshold values of 5 s for TTC and 4 s for PET are adopted in this study. Following the recommendations of Shahdah et al. [50], 60 simulations per scenario are conducted with these thresholds, limiting the error to 5%. The SSAM is then utilised to extract conflict profiles below these thresholds from the simulated trajectory data, with parameters such as conflict angle, change in velocity, and minimum time to collision further investigated to assess potential crash risks. Conflicts with high likelihood and significant consequences are penalised in the calculation of safety costs. The total safety cost representing a particular scenario are subsequently calculated by averaging the real-time safety performance across 60 simulations, each lasting 15 min, for each set of traffic conditions.

Building on the conflict data extraction, this study presents a cost-based safety evaluation pipeline that considers both the probability and severity of the potential events. The model outlines the process for assessing safety costs based on the conflict profiles derived from the SSAM. This includes evaluating the "consequences" and "likelihood" of near-miss events, as well as converting safety risk into cost values by accounting for the economic and human costs associated with potential consequences.

3.1.1. Consequence of crash

Various studies have explored the relationship between crash severity and risk factors, including speed, age, and seatbelt use [18,56–60]. The World Health Organisation's global road safety report [61] identifies speed as a major cause of fatal and severe injuries (FSI). Additionally, change in velocity (delta-V) during a crash is commonly used as a quantitative measure of severity. Joksch [58] concluded a rule of thumb proportional relationship between fatality risk in vehicular crash and the fourth power of delta-V. More recently, Bahouth et al. [56], Brumbelow [57] and Shannon et al. [60] provided further proof of the significant role of the delta-V using a complex model, which can differentiate between crash configurations included head-on and rear-end collisions.

This study adopts the model developed by Bahouth et al. [56] to derive the consequence and delta-V relationship for all the four types of crashes, namely head-on, near side, far side and rear-end (Fig. 4). According to the author, probability of a driver suffers MAIS3+ injury is denoted using a binary logistic regression model, mathematically expressed as

$$Pr(FSI) = \frac{1}{1+e^{-t}}, \tag{6}$$

where

$$t = \beta_0 + \beta_1(deltaV) + \beta_2(AACN) + \beta_3(belt_{use}) + \beta_4(Age55-75) + \beta_5(Age75+). \tag{7}$$

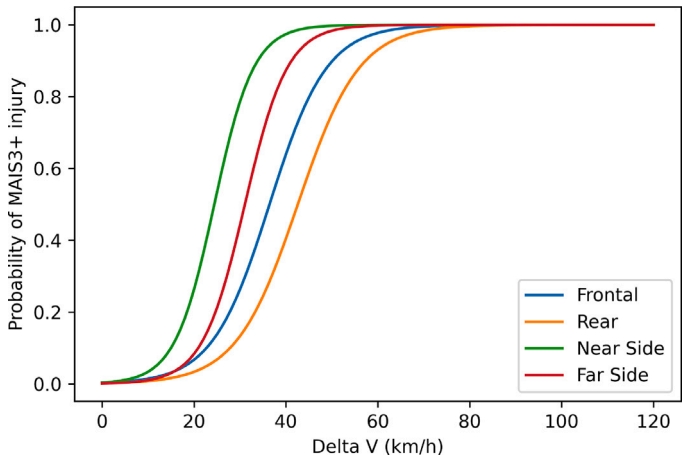

Fig. 4. Illustration of consequence of crash vs. delta-V.

Table 3
Coefficient for crash configurations.

Conflict type	$t = \alpha_0 \times \text{deltaV} - \alpha_1$	
	α_0	α_1
Frontal	0.1604	5.8446
Rear	0.1492	6.3549
Near side	0.2333	5.6989
Far side	0.2167	6.735

In the right-hand of Eq. (7), β is the coefficient to be estimated corresponds to crash configuration and predictors except delta-V are Boolean values. For instance, if the impact is severe enough to trigger the Advanced Automatic Collision Notification (AACN) system, then the AACN would be equal to 1 (True). By assuming driver age between 16–54, Seat belt is always used and the AACN is triggered, expression of t for the 4 crash configuration will be simplified as in Table 3.

On the other hand, crash configuration can be determined using conflict angle [51] (Fig. 5). Thus, probability of a driver suffering a fatal and severe injury (FSI) is described as a function of the conflict angle and delta-V. Considering only a two-vehicle accident, the probability of any one of the two driver suffering FSI can be denoted as the union of the probabilities of individual driver suffering FSI, as formulated by

$$Pr(FSI_1 \cup FSI_2) = Pr(FSI_1) + Pr(FSI_2) - Pr(FSI_1) \times Pr(FSI_2). \tag{8}$$

3.1.2. Likelihood of FSI crash

With regard to likelihood of crash, considerable amount of research has been conducted to explore the connection between conflict indicators and crash occurrence [20,22,23,62–65]. These studies attempted to estimate crash potential using conflict indicators and analytic methods, such as EVT. However, there are still concerns about the inconsistency between conflicts and crashes, as the information provided in the conflict data are often limited to the conflict type, choice of threshold, and conflict data collection date and time, location and technology used [65]. For this study, given the complexity of the research question on crash potential, the approach presented by Oh and Kim [64] is utilised which uses single measure of crash nearness, TTC. According to the author, probability of crash can be denoted as

$$Pr(C) = a + b \times e^{\frac{-ttc}{c}}. \tag{9}$$

From Eq. (9), let coefficient a and b to be 0 and 1, respectively as probability of crash ranges from 0 to 1. Values of coefficient c affects the rate of decay (Fig. 6), which depends on weather condition, driver reaction time and age, etc. Assuming average driver reaction time of 1.5 s, $c = 0.5$ is selected and the relationship becomes

$$Pr(C) = e^{\frac{-ttc}{0.5}}. \tag{10}$$

Hence, probability of FSI crash from a near-miss event can be evaluated as the product of consequence and likelihood, that is,

$$Pr(FSI and C) = Pr(FSI_1 \cup FSI_2) \times Pr(C). \tag{11}$$

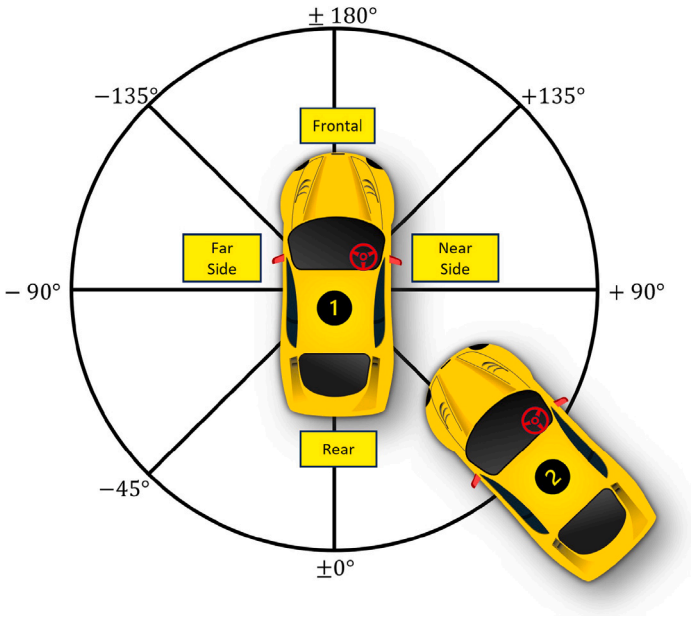

Fig. 5. Schematic of conflict angle and conflict configuration.

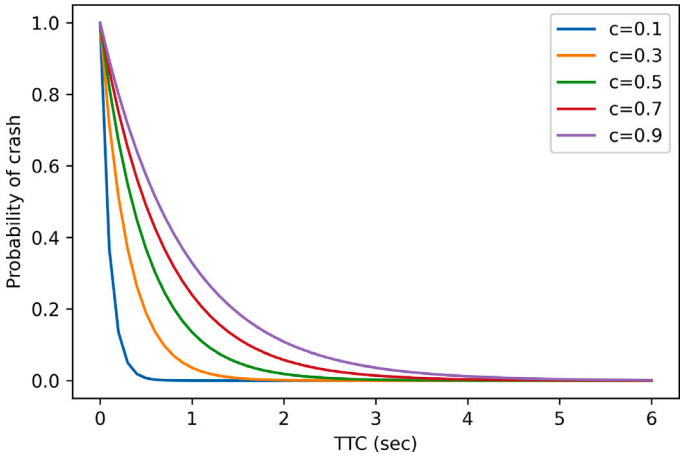

Fig. 6. Illustration of crash potential vs. time to collision at different rate of decay.

Thus, it is a function of conflict angle, delta-V and TTC, equivalently,

$$Pr(FSIandC) = f(conflictangle, deltaV) \times g(ttc). \tag{12}$$

3.1.3. Convert safety risk into cost

To ensure comparability between safety and other objectives, such as efficiency and sustainability, within the reinforcement learning (RL) framework, safety risk is converted into an equivalent monetary cost. Previous research has employed several methods to quantify crash events in monetary terms, including the human capital (HC) approach and the value of risk reduction (VRR) method [66–69]. The HC approach accounts for direct and identifiable economic losses such as hospitalisation expenses, loss of income, and property damage. In contrast, the VRR method reflects societal preferences through individuals' willingness to pay (WTP) to avoid injury or fatal crashes. The studies by Hensher et al. [67] and guidelines from Transport for New South Wales [69] offer robust methodologies for estimating WTP. Notably, Transport for New South Wales [69] advocates for an inclusive WTP approach consistent with recommendations from the Australian Government Department of Infrastructure, Transport, Cities and Regional Development (DITCRD) and adopted by the Australian Transport Assessment and Planning (ATAP) Guidelines. Typically, these monetisation approaches have been applied primarily within planning domains to support long term investment decisions related to infrastructure upgrades and traffic signal design.

Table 4
Inclusive willingness-to-pay per accident.
Source: Reprinted from Transport for NSW Economic Parameter Values (p.31), 2022, Transport for New South Wales, TfNSW.

Severity level	Urban	Rural	Average
Fatality	$7,808,768	$9,242,523	$8,586,767
Serious injury	$507,553	$700,151	$574,265
Moderate injury side	$85,296	$112,608	$97,512
Minor injury side	$78,389	$103,484	$89,314
Property damage only	$10,338	$10,338	$10,338

In contrast to such traditional retrospective planning applications, the method proposed in this study applies monetary risk conversion in real time operations. Instantaneous near miss events detected within each signal cycle are dynamically quantified into monetary cost metrics based on both their likelihood and consequence. These metrics are integrated directly into the deep RL optimisation process, allowing immediate and continuous adaptation to changing traffic conditions. Although our estimation techniques align closely with established WTP methodologies recommended in existing literature, the novelty of our approach lies explicitly in its real time operational deployment. This enables the RL algorithm to concurrently optimise safety and efficiency at an operational level, rather than solely informing long term planning or infrastructure investment decisions. The inclusive WTP values recommended by Transport for New South Wales and applied in this study are detailed in Table 4.

In this study, the focus is on urban intersections, where the safety cost of a near-miss event can be calculated through a two-step process using the urban inclusive WTP. Initially, the consequence of a potential crash is mapped to its corresponding severity level to determine the maximum potential cost, mathematically expressed as

$$maxWTP(x) = \begin{cases} 7808768 & \text{if } Pr_{FSI}(x) \geq 0.98 \\ 507553 & \text{if } 0.7 \leq Pr_{FSI}(x) < 0.98 \\ 85296 & \text{if } 0.45 \leq Pr_{FSI}(x) < 0.7 \\ 78389 & \text{if } 0.2 \leq Pr_{FSI}(x) < 0.45 \\ 10338 & \text{otherwise.} \end{cases} \tag{13}$$

This cost is then multiplied by the probability of a fatal crash to estimate the safety cost associated with the conflict. Considering all conflicts, $x \in D$, happened during a specific period, the total safety cost is evaluated as

$$WTP_{total} = \sum_{x \in D} Pr_{crash}(x) \times Pr_{FSI}(x) \times maxWTP(x). \tag{14}$$

3.2. RECOSAM architecture

The use of various machine learning algorithms in traffic safety analysis has been documented in prior research [18,26,39,70,71]. In this study, regression evaluation model is developed and assessed using several machine learning algorithms, including extreme Gradient Boosting (XGBoost), Random Forest (RF), and Multi-Layer Perceptron (MLP). Among these, XGBoost was ultimately adopted as the core model for the proposed ensemble framework due to its superior performance.

XGBoost is an ensemble machine-learning algorithm that combines multiple decision trees to improve predictive accuracy. It is widely recognised as a scalable tree boosting system and has demonstrated success across various domains, such as hazard risk prediction and retail sales forecasting [72]. All models, including RF and MLP, were implemented using the `scikit-learn` package in Python [73].

To identify optimal hyperparameters, a combination of grid search and 10-fold cross-validation was employed. This approach evaluates model performance across unseen samples, where each fold serves as a validation set while the remainder act as training data. Compared to a conventional single train–test split, this method reduces bias and enhances generalisability. Key hyperparameters considered during tuning included the learning rate, maximum tree depth, and the number of estimators. Model performance was assessed using the root mean square error (RMSE) and the coefficient of determination (R^2).

The architecture of RECOSAM, depicted in Fig. 1 path (1), consists of seven sub-models aligned with the phase configurations in Table 1. Comparative results of the ensemble framework using XGBoost, RF, and MLP are presented in Table 5. The XGBoost-based ensemble consistently achieved higher R^2 values on average, with performance improvements ranging from 0.52 to 0.69. While the MLP model exhibited comparable accuracy in some configurations, the XGBoost model outperformed both alternatives, justifying its selection for the final model architecture.

Table 2 outlines the input variables for RECOSAM, categorised into traffic demand and signal action variables. Traffic demand variables represent observable features used by ATSC algorithm for operational decisions. Each query to the RECOSAM includes traffic demand variables from both intersection sides resulting in six variables. Signal action variables refer to responsive signal control operations, with each query assessing possible operations and returning near-future safety costs. With phase types limited to two operations, the traffic signal associated inputs comprises variables, including the two elementary phase's duration, waiting time due to the previous stage, and the phase type to be executed. RECOSAM is designed to estimate forthcoming safety risks for

Table 5
Performance of initial and proposed ensemble framework.

Model	Coefficient of determination (R^2)						
Initial evaluation model	0.21						
Ensemble evaluation model	Permissive	LeadPR	LagPR	PRPT	PTPR	LeadPT	LagPT
XGBoost	0.61	0.52	0.64	0.64	0.69	0.52	0.65
RF	0.57	0.52	0.56	0.56	0.49	0.40	0.50
MLP	0.63	0.55	0.57	0.57	0.55	0.49	0.59

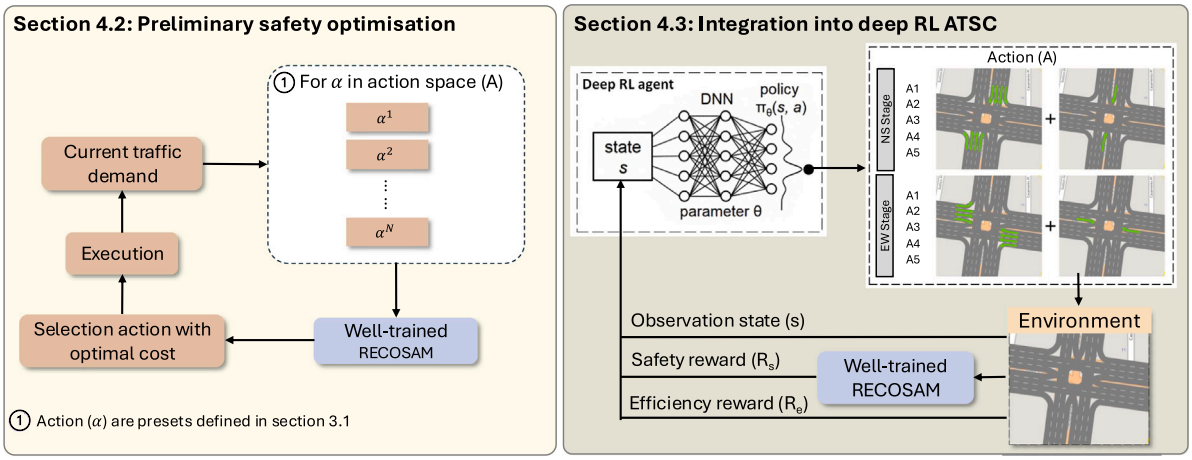

Fig. 7. RECOSAM applications.

each query rather than offering optimal safety configurations directly. This align with the use case in the deep RL algorithm, which involved a trial-and-error learning process, balancing safety and efficiency within reasonable time frames. By leveraging this model, deep RL agents can weigh safety and efficiency objectives in their decision-making, promoting road user safety while optimising efficiency. Subsequent sections will explore its capabilities across various situations.

4. Experiments and case study

This section presents two experiments and a case study. Firstly, we assess model variability under specific demand scenarios. Next, we implement RECOSAM in a preliminary safety optimisation algorithm, where it evaluates all possible actions to inform operational decisions. Finally, we demonstrate the integration of RECOSAM into a deep RL ATSC, utilising it as the safety cost calculation component. The real-time safety evaluation provided by RECOSAM enhances the efficiency of training deep RL models.

RECOSAM is meant for integration into ATSC algorithms, with two potential use cases illustrated in Fig. 7. The details of the corresponding experiments and case study are provided in Sections 4.2 and 4.3, respectively. It is important to note that this study utilised the same simulation model throughout the data synthesis and ATSC use cases.

4.1. RECOSAM variability

Following sections explore the model variability under the through dominate demand scenarios, as detailed in Table 6. The scenarios represent traffic demand during off-peak hour, peak hour and overstated peak hour. The results were graphically presented, depicting the following relationship:

1. Safety cost per second against waiting time ($W_{\lambda,s-1}$) and action time (T_s).
2. Safety cost per second against ratio of the first elementary phase duration (D_1) to the action time.

In terms of safety risk evaluation, the first relationship pertains to the potential safety risk that results from both uncontrollable and controllable factors, such as waiting time (red light) and upcoming action (green light) time, respectively. This relationship allows for a comparison of phase configuration across different demand scenarios. On the other hand, the second relationship pertains to the potential safety risk associated with the internal settings of each phase configuration. In these plots, the larger value on horizontal axis (maximum in 1) indicates the first elementary phase occupies a greater proportion of action time than the second one. Since the Permissive only has a single phase, it appears as a vertical pattern. This relationship provides insight into the effect of elementary phase on the internal phase configurations.

Table 6
Evaluation scenarios.

Scenario No.	Traffic demand		Turning demand		
	Definition	veh/h	Left turn %	Through %	Right turn %
1	Oversaturated peak hour	400	20	60	20
2	Peak hour	250	20	60	20
3	Off peak hour	100	20	60	20

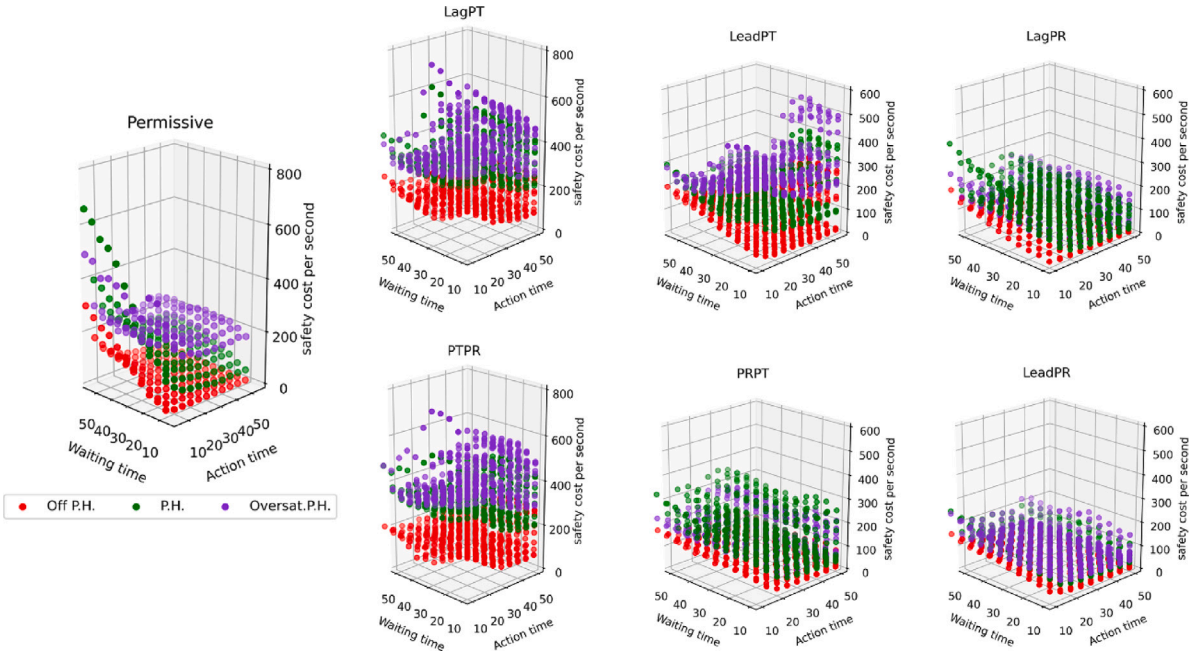

Fig. 8. Type 1 plots for all phase options with various traffic demands.

The figures depict results for a scenario where through traffic flow dominates, comprising 60% for through, 20% for right turns, and 20% for left turns at the given intersection. This scenario typically arises at intersections where a major road intersects with a minor road. As shown in Fig. 8, it is worth noting that increasing the action time may lead to higher safety risks in certain cases, such as LeadPT, LagPT, and PRPT. This implies that extending the green duration may not be the optimal choice in through-dominant situations, particularly when the waiting time due to previous stages is short. Moreover, during peak-hour demand, LagPR and PRPT exhibit higher safety risks compared to situations of oversaturated demand.

Regarding the internal configuration, Fig. 9 suggests that having approximately equal lengths of elementary phases ($D_1 \approx D_2$) in LagPT, PTPR, and LeadPT configurations may lead to increased safety risks, particularly in scenarios with long waiting times or oversaturated traffic volumes. Conversely, the LagPR configuration demonstrates a positive correlation, whereas LeadPR and PRPT exhibit a negative correlation, forming a crossing pattern. This pattern indicates that increasing the proportion of dedicated right-turn phases can mitigate risky conflicts and enhance road user safety.

The findings indicate that safety risk increases with action time and decreases with waiting time. This aligns with traffic conflicts primarily arising from car-following, especially during stop-and-go interactions. Therefore, keeping vehicles stationary may serve as a safety-biased solution to mitigate risks. However, this strategy conflicts with efficiency optimisation strategies, which aim to maximise action time and minimise waiting time, emphasising longer green phases for major traffic flows to maintain continuous movement. Moreover, while increasing the proportion of dedicated right-turn phases can enhance safety, excessive prioritisation of these phases may lead to inefficient green time allocation, thereby restricting the balanced and efficient flow of overall traffic.

4.2. Preliminary safety optimisation experiments

RECOSAM is tested with preliminary optimisation experiments to assess safety costs under diverse traffic conditions. The model determined the optimal phase configuration based on traffic demand information. Results from two scenarios, representing an arbitrary hour, are presented. The optimisation workflow, shown in Fig. 7 application (1), employs a trial-and-error method. It pairs demand variations with randomly selected actions from the signal action space (Table 1). In order to demonstrate the optimisation problem, only the North–South (NS) approach is considered, assuming steady demand in the East–West (EW) direction. Fixed-time

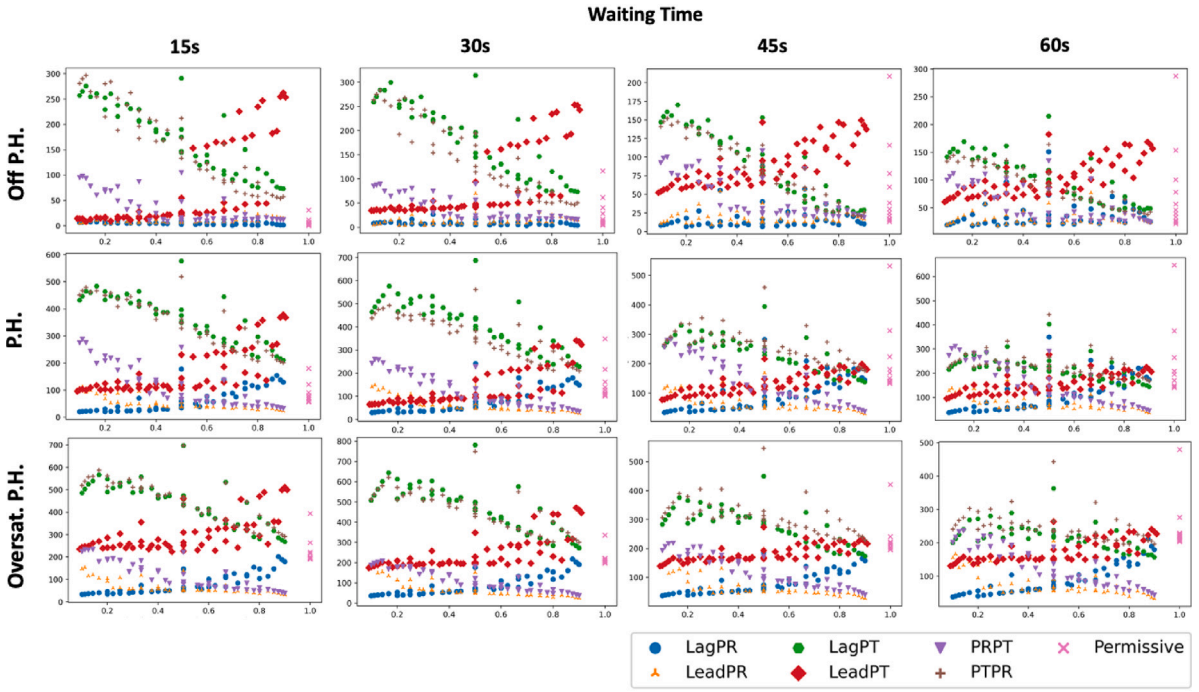

Fig. 9. Type 2 plots for various traffic demands and waiting time. Each subplot represents the safety cost per second as a function of the ratio between the first elementary phase duration and the total action time (D_1/T_s).

traffic signals are used for the East–West approaches. These experiments demonstrate the utility of RECOSAM and provide insights into selecting optimal signal configurations.

4.2.1. Scenario 1

The first scenario is focused on a situation with a predominant demand for turning movements, simulating an increasing proportion of traffic volume making turns at the intersection ($\theta_{\lambda,i}$) for specific destinations, like schools or large parking facilities. Left turn volume for northbound traffic and right turn volume for southbound traffic gradually increased (Table 7), creating induced demand towards the westbound approach. Traffic volumes for eastbound and westbound movements remained steady. The traffic signal for east–west approaches assumed a 20-s fixed-time signal.

Fig. 10 and Table 8 present the optimal average cost and the corresponding phase configurations for each quarter-hour period, respectively. It is observed that RECOSAM suggested relatively short permissive phase configurations when the traffic demand is low. However, as the demand and turning proportions grow, RECOSAM tends to extend phase durations and utilise configurations that provide exclusive protection to the conflicting flow. This is evident from the internal elementary phase proportion. Taking the optimal configuration of the 3rd and 4th quarters as an example, LeadPR has a protected right phase followed by a permissive phase (Table 1). The proportion of the protected right turn phase to the total phase duration increased with the increasing turning volumes.

The difference between LagPR (Lagging Protected Right) and LeadPR (Leading Protected Right) lies in the arrangement of the protected right phase. The trend of selecting LeadPR suggests that arranging the protected right turn phase at the beginning of phases could help discharge the turning demand in a safer manner. On the other hand, the optimal average cost increased rapidly with respect to increased total traffic volumes. This is reflected in the average cost and the cost per green time allocated. This could be because there is a limitation on action time (T_s), with the maximum being 60 s in total, including transition time (T_t), to prevent overextended phase duration. Consequently, safety risks could not be reduced through an further extension of phase length when traffic started to saturate.

4.2.2. Scenario 2

The second scenario explores traffic conditions characterised by fluctuating demand on a typical weekday with moderate traffic. The demand exhibits oscillations without a discernible trend in total volume but manifests variations in the proportion of turning movements (Table 9). Traffic volumes for eastbound and westbound movements remained steady. In the context of this scenario, the eastbound and westbound traffic volumes remain steady at an average of 60, 45, and 45 vehicles per quarter-hour for through, right turn and left turn movements, respectively. A 30 s fixed-time signal is assumed for the east–west approaches.

Table 7
Scenario 1 demand variation.

Direction	00:00–15:00	15:00–30:00	30:00–45:00	45:00–60:00
Total volume				
Northbound	100	150	200	250
Eastbound	100	100	100	100
Southbound	100	150	200	250
Westbound	100	100	100	100
Right turning volume				
Northbound	20	15	20	25
Eastbound	20	20	20	20
Southbound	20	90	120	175
Westbound	20	20	20	20
Left turning volume				
Northbound	20	90	120	175
Eastbound	20	20	20	20
Southbound	20	15	20	25
Westbound	20	20	20	20
Through movement volume				
Northbound	60	45	60	50
Eastbound	60	60	60	60
Southbound	60	45	60	50
Westbound	60	60	60	60

Table 8
Scenario 1 optimal signal configuration.

Period	Phase (P)	Elementary duration 1 (D_1)	Elementary duration 2 (D_2)	Total
00:00–15:00	Permissive	15	N/A	15
15:00–30:00	LagPR	45	5	50
30:00–45:00	LeadPR	20	30	50
45:00–60:00	LeadPR	45	5	50

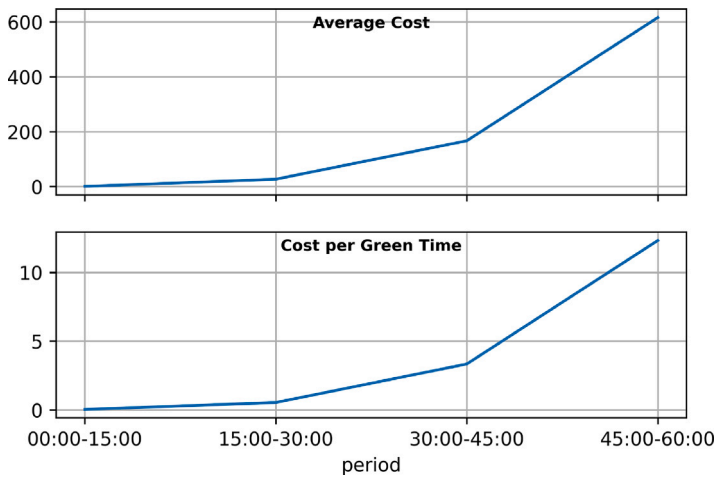

Fig. 10. Scenario 1 optimal average cost.

Fig. 11 and Table 10 display the optimal average cost and corresponding phase configurations for each quarter-hour period. With an increase in demand from 150 to 200 vehicles per quarter-hour, the optimal phase configurations shift from LeadPR to PTPR. This suggests adaptive behaviour to manage traffic flow safely, such as prolonging the protected through phase in the 3rd quarter-hour may be a strategic response to accommodate higher left turn volumes.

The optimal cost per green time increases with higher traffic volumes, as expected, since more complex traffic situations generally incur higher costs. However, a comparison between the 3rd and 4th quarter-hours reveals a significantly higher total cost during the latter period. This suggests that RECOSAM tends to allocate more green time to mitigate safety risks as demand rises, which may lead to an unfair and inefficient distribution of right of way. This finding underscores the critical importance of balancing safety and efficiency in advanced traffic control systems.

Table 9
Scenario 2 demand variation.

Direction	00:00–15:00	15:00–30:00	30:00–45:00	45:00–60:00
Total volume				
Northbound	150	125	200	175
Eastbound	150	150	150	150
Southbound	150	125	200	175
Westbound	150	150	150	150
Right turning volume				
Northbound	45	50	40	53
Eastbound	45	45	45	45
Southbound	45	50	40	53
Westbound	45	45	45	45
Left turning volume				
Northbound	52	25	120	53
Eastbound	45	45	45	45
Southbound	52	25	120	53
Westbound	45	45	45	45
Through movement volume				
Northbound	53	50	40	69
Eastbound	60	60	60	60
Southbound	53	50	40	69
Westbound	60	60	60	60

Table 10
Scenario 2 optimal signal configuration.

Period	Phase (P)	Elementary duration 1 (D_1)	Elementary duration 2 (D_2)	Total
00:00–15:00	LeadPR	15	35	50
15:00–30:00	LagPR	40	10	50
30:00–45:00	PTPR	25	5	30
45:00–60:00	LeadPR	15	35	50

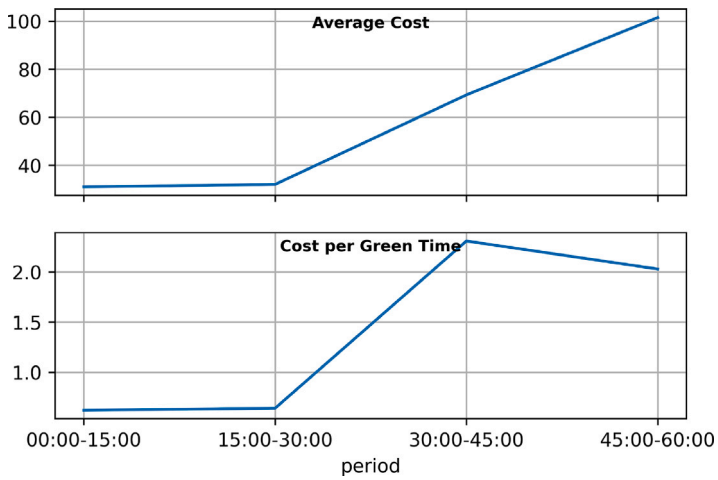

Fig. 11. Scenario 2 optimal average cost.

In summary, these results showcases the adaptability of RECOSAM to diverse traffic scenarios. The model makes specific adjustments in terms of phase type selection and the duration of its elementary phases. Offering a shortcut alternative to the time-consuming simulation and offline evaluation processes. The model demonstrates its utility in trial-and-error optimisation problems, such as those encountered in deep reinforcement learning frameworks.

4.3. Safe deep reinforcement learning ATSC case study

A case study has been conducted within the AIMES testbed, established by the University of Melbourne in 2016. The study focuses on the Elgin-Rathdowne intersection simulation model (Fig. 12), which was also employed in developing RECOSAM. The

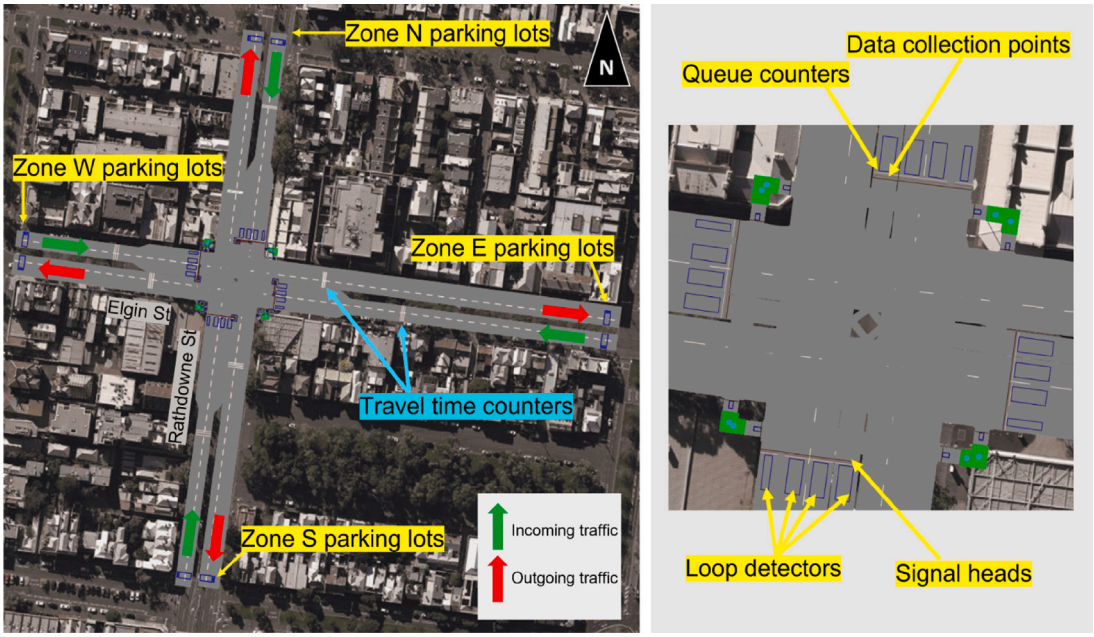

Fig. 12. Elgin-Rathdowne intersection simulation model. *Note:* From "Multi-Modal Traffic Signal Control using Deep Reinforcement Learning", by Yazdani [74], The University of Melbourne. Reprinted with permission.

subsequent section presents the case study set up, followed by presentation of results and discussion. It is noted that the case study is to demonstrate an potential application of RECOSAM. The action explored is only a subset of the action spaces available in RECOSAM.

4.3.1. Integrating RECOSAM into deep RL traffic signal

RECOSAM serves as an edge computation tool integrated into the existing deep RL-based traffic signal control framework. Known as Safe-RL, this integration introduces a safety cost component alongside efficiency considerations during decision-making, as illustrated in Fig. 7 application (2). The decision-making process begins by observing the current environment, incorporating metrics like safety costs and delays to select the optimal signal configuration. Subsequently, the chosen configuration is evaluated based on safety and efficiency performances, refining the learning process iteratively.

The deep RL model learns through trial and error until it converges on the optimal choice. However, the time-intensive safety assessment pipeline shown in Fig. 1 path (2) is impractical. This limitation creates an opportunity to develop RECOSAM, which provides a shortcut for estimating safety costs during decision-making. Consequently, both the trial-and error learning process and decision time in real-time optimisation are reduced.

Concurrently, ensuring comparable units for rewards is crucial for the deep RL agent to balance safety and efficiency effectively. Hence, the safety evaluation pipeline converts safety risks into a monetary cost equivalent, facilitating meaningful comparisons with economic losses due to traffic inefficiency. For the Safe-RL model presented in this paper, a delay cost per vehicle hour of $36.3 [69] is applied to convert efficiency loss into the monetary cost equivalent.

4.3.2. Signal model settings

The study aims to experiment with the integration of RECOSAM into the deep RL based traffic signal control proposed by Yazdani et al. [6]. In the following sections, the original deep RL based traffic signal control, which solely aimed at optimising efficiency will be referred to as the Baseline RL. On the other hand, the model that considers both safety and efficiency in the decision-making procedure will be referred to as Safe RL, discussed in Section 4.3.1. The performance of the Safe RL based traffic signal control is compared against the Baseline RL model in terms of overall safety and efficiency performance measures (PMs).

In the experimental setup, each model undergoes training for 1200 episodes, with each episode corresponding to a 1-h simulation period. To ensure reproducibility, each model is replicated using 4 different random seeds. The simulation environment runs with different random seeds to account for traffic arrival patterns' stochasticity. There are 10 actions for the agent to be selected. These actions include a permissive phase followed by a protected right phase, terminated by a transition to amber and red signals. Action settings are detailed in Table 11.

In this study, demand variations collected from the real-world location, representing an hour-long period during an arbitrary weekday, were utilised. Table 12 presents the demand variation settings over the 1-h simulation. It is assumed that the demand in the opposite direction is similar, for instance, north-bound and south-bound traffic will have the same level of demand.

Table 11
Safe RL action for each active stage.

Stage	Action configuration (s)				
	No.	Permissive	Protected right	Transition	Total
NS	1	10	5	5	20
	2	15	5	5	25
	3	20	5	5	30
	4	25	5	5	35
	5	30	5	5	40
	No.	Permissive	Protected right	Transition	Total
EW	1	10	5	5	20
	2	15	5	5	25
	3	20	5	5	30
	4	25	5	5	35
	5	30	5	5	40

Table 12
Demand variation (PCU) for 1 h simulation.

Direction	00:00–15:00	15:00–30:00	30:00–45:00	45:00–60:00
Total volume				
NS	300	400	350	300
EW	250	350	300	250
Right turning volume				
NS	30	40	35	30
EW	25	35	30	25
Left turning volume				
NS	30	40	35	30
EW	25	35	30	25
Through movement volume				
NS	240	320	280	240
EW	200	280	240	200

Table 13
Performance metrics results of last 50 training episodes (Trained model).

Model	Performance measures (PMs)				
	Cumulative costs ($)			Queue length (m)	Travel time (s)
	Safety	Delay	Delay + Safety		
Baseline RL	1375.6	966.44	2342.04	23.72	44.94
Safe RL	1356.39	1024.03	2380.42	24.74	46.76
Severity RL	210 375.53	3641.42	214 016.95	55.30	82.15

4.3.3. Result and discussion

The experimental results are presented in two main groups, based on different performance measures (PMs), as follow:

1. Model performance throughout training episodes. This group includes PMs such as cumulative delay cost, cumulative safety cost, average queue length, and average travel time.
2. The policy analysis of models. In this group, we examine the selection frequency of deep RL actions and green time allocated to stages.

These experimental results are depicted in qualitative (plots) and quantitative (table) representation, reporting both the mean values and standard deviation. The qualitative results depict the model's average performance throughout the training episodes using a shaded area. Furthermore, to gain insight into performance changes before and after integrating RECOSAM, the quantitative results compare the model's performance of the first 50 episodes with 95% confidence interval.

Fig. 13 illustrates the training performance of the models over 1200 episodes. As shown, the Safe RL model (green) surpasses the Baseline RL model (blue) in terms of cumulative safety costs after approximately 600 episodes. This observation indicates that incorporating safety considerations into the deep RL reward function can enhance overall safety performance.

However, this improvement comes at the expense of reduced efficiency, as evidenced by slightly higher cumulative delay costs, average queue lengths, and average travel times after a sufficient number of training episodes. Nevertheless, the sacrifice in efficiency

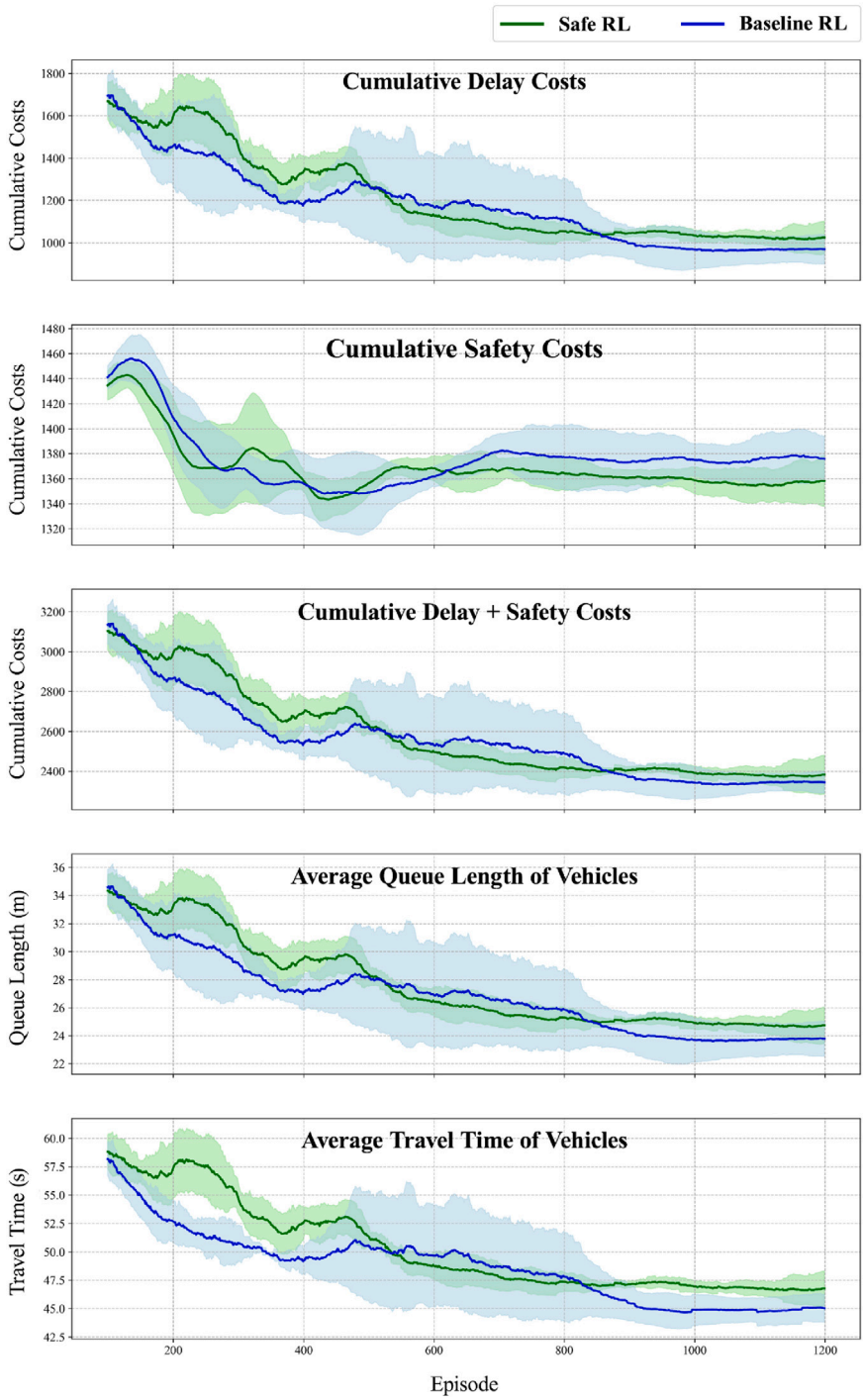

Fig. 13. Model Performance throughout episodes.

is relatively modest. According to the average performance results presented in Table 13, the Safe RL model improves safety-associated costs by an average of 1.42% over the last 50 training episodes. In contrast, delay costs, queue lengths, and travel times increase by 5.62%, 4.12%, and 3.89%, respectively. Overall, the total cumulative cost rises by 1.61%.

Although the cumulative cost performance of the Safe RL model appears 1.16% worse than that of the Baseline RL model, it is important to note that there are limitations on the available action sets. In other words, the full capabilities of RECOSAM were not fully exploited. Furthermore, the Baseline RL represents a single-objective optimisation focused solely on efficiency, whereas

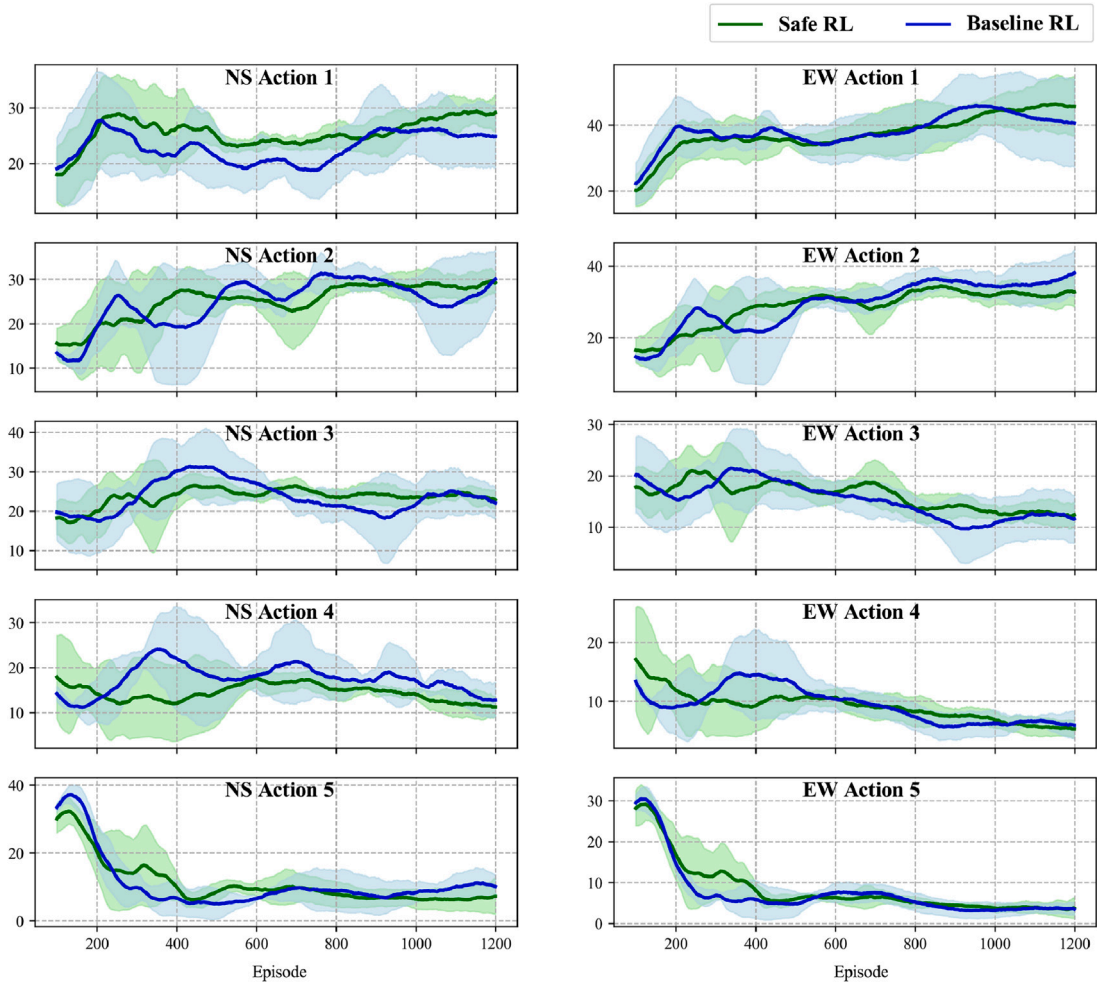

Fig. 14. Action selection policy of the RL models during the 1200 training episodes. Each subplot corresponds to the actions listed in Table 11. *Y*-axis represents the proportion of actions selected (in %).

the Safe RL represents a multi-objective optimisation balancing both safety and efficiency. This inherently introduces challenges in achieving significant improvements across all objectives simultaneously. Nevertheless, the presented Safe RL model demonstrates the potential to manage these two critical objectives with minimal trade-offs.

Additionally, the performance of the Severity RL model is also reported in Table 13. This model represents a scenario in which the likelihood component of crash risk is inhibited in RECOSAM, meaning that the safety evaluation considers only the severity of conflicts, assuming that each conflict results in a crash. The results clearly show that the Severity RL model introduced a strong safety bias. The estimated safety cost ($210,375) overwhelmingly dominated the delay cost ($3,641), leading to substantial increases in both the average queue length and travel time compared to both the Baseline RL and Safe RL models. These results demonstrate that a safety-biased optimisation is not favourable for practical multi-objective control and highlight the importance of RECOSAM's capability to consider both crash likelihood and severity, combining them into a unified framework for multi-objective optimisation.

Fig. 14 visualise the action selection policy of the Safe RL during training episodes. The results reveal that the Safe RL tends to select actions with shorter action time (T_s). In contrast to the Baseline RL, which allocates more green time to the NS stage (major approaches), the Safe RL shortens the duration of each stage (both NS and EW) to reduce safety risks. Longer green time in one stage leads to extended waiting times in the competing stage, which has been identified as a factor contributing to increased safety costs in previous section. Hence the Safe RL opts for shorter actions in both NS and EW directions to mitigate potential safety risks, resulting in shorter signal cycles and a more balanced allocation of green time.

Similarly, Fig. 15 illustrates the same phenomenon. The action policy pie charts indicate that the Safe RL approach achieves a more balanced allocation of green time between the different directions, with 53.8% allocated to NS and 46.2% to EW, compared to the Baseline RL, which allocates 54.9% and 45.1%, respectively. Examining the distribution of selected actions within the NS direction (major approaches), it becomes evident that the Baseline RL tends to favour longer durations (T_s), with 14.7% of action 4 and 10% of action 5 being selected. In contrast, Safe RL reduces the proportion of action 4 by 2.7% and action 5 by 3.4%. Conversely,

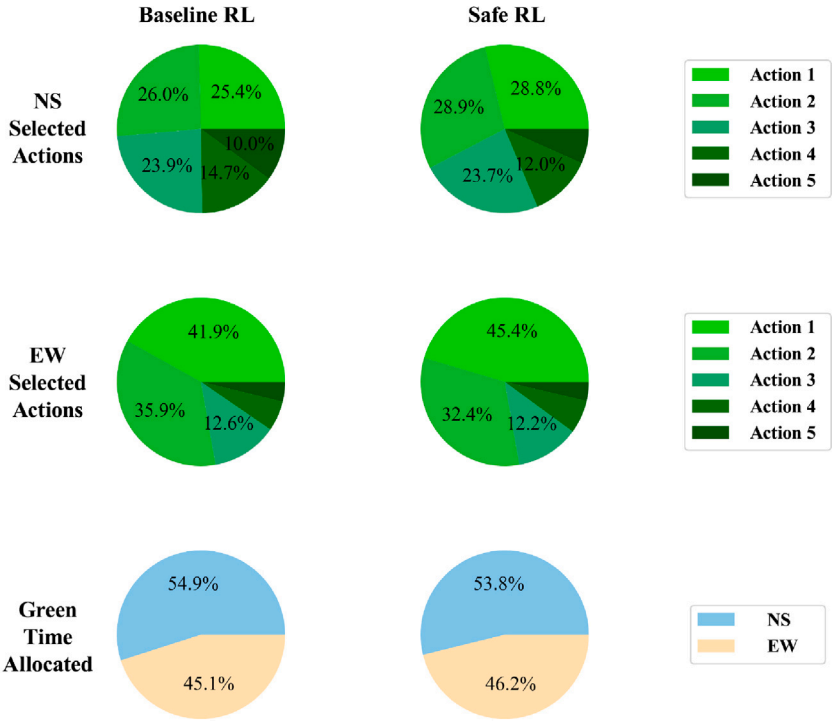

Fig. 15. Action policy pie charts (last 50 episodes).

Safe RL increases the proportion of action 1 by 3.4% and action 2 by 2.9%. This suggests that Safe RL prefers shorter action time for NS stages when they are selected.

For the EW direction (minor approaches), Safe RL also increases the proportion of action 1 from 41.9% (Baseline RL) to 45.4%. The increased proportion of action 1 for both directions in Safe RL results in a reduction of average action time, leading to shorter and more frequent signal cycles. It is important to note that as the proportion of green time allocated to EW stages increases, the Safe RL's optimal policy priorities minor approaches. However, this policy preference for shorter actions leads to more frequent signal operations overall.

5. Conclusion

In this paper, a real-time cost-based safety prediction model (RECOSAM) is developed and tested for preemptive safety risk prevention in advanced traffic signal control systems, such as deep reinforcement-based frameworks. RECOSAM evaluates safety risk by incorporating crash likelihood and crash consequences as an expected risk converted to equivalent costs, providing a justifiable and comparable measure alongside other optimisation objectives, such as queue length reduction or travel time saving.

RECOSAM is tested across various traffic volume scenarios, mainly dominated by through traffic demand volumes. Results reveal two key relationships: safety risk linked to uncontrollable and controllable factors, and safety risk associated with internal phase configurations.

On a more general level, findings show safety risk correlates positively with action time (for the direction under study) and negatively with waiting time (time utilised for cross direction), potentially conflicting with efficiency optimisation. More specific to conflicts associated with right turns (in left-side driving), increasing "dedicated right turn" phases reduces safety risk, but excessive extension may hinder traffic flow efficiency. This is where the RECOSAM shows value, namely in finding the optimal trade-offs between efficiency and safety. For example, in a basic peak/offpeak traffic situation, RECOSAM will most likely recommend dedicated right turn phases to mitigate high risks, while permissive right turns during off-peak to maintain efficiency without compromising safety. The advantage of RECOSAM is to make this decision dynamically based on observed demand volumes in real-time, which is much more valuable especially in central districts with demand density or busy major arterial which may not display the typical peak/offpeak demand pattern throughout the day.

RECOSAM was also tested with preliminary optimisation experiments to investigate its capability in assessing safety costs for various traffic conditions, and reflecting risk costs in traffic signal control decisions. Two prevalent traffic scenarios were examined, and the results demonstrated the effectiveness and flexibility of RECOSAM. The model makes precise adjustments regarding phase configuration selection and the duration of its elementary phases. This experiment demonstrates that RECOSAM is a viable and effective alternative to the offline evaluation procedures, and showcases its efficacy in real-time optimisation challenges.

Lastly, RECOSAM is integrated into a deep RL framework and undergoes evaluation in a real intersection case study. The primary findings suggest that integrating safety considerations has the potential to enhance the safety of the traffic environment. However, it is observed that the delay costs outweighs the associated safety costs. This observation is consistent with previous evaluation outcomes.

In terms of policy implications, the proposed safety model can offer significant benefits for operating intersections with historically higher rates of crashes or accidents, or zone with higher population of elderly drivers/residents or school/hospital zones. Such selective approach has the potential to enhance safety at targeted intersections while maintaining efficiency at others. Moreover, prioritising safety considerations may be particularly crucial at intersections operating within multimodal environments, including vulnerable users such as cyclists and pedestrians.

6. Limitations and future directions

Although the overall performance of Safe RL appears to be inferior to that of the Baseline RL, there are limitations on the available actions. In other words, the full capabilities of RECOSAM were not examined. Additionally, Baseline RL is a single objective optimisation focused solely on efficiency, whereas Safe RL is multi-objective, making it challenging to optimise both goals simultaneously. Despite these challenges, Safe RL shows potential in balancing safety and efficiency with minimal trade-offs. Further research is needed to validate the comparability of safety and delay costs.

This paper introduces RECOSAM for ML-based ATSC aimed at preemptive risk prevention. The model provides near-future safety costs for deep RL to incorporate into traffic operation decisions. Its key advantage is real-time applicability, enabling proactive risk management. While current study focuses on isolated intersections and vehicle traffic, future work will expand the methodology to multimodal environments such as including pedestrians and public transport, as well as to the network of intersections.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is funded by Australian Research Council (ARC) LP200301389, Kapsch TrafficCom Australia, Royal Automobile Club of Queensland (RACQ), and iMOVE CRC, the Cooperative Research Centres program, an Australian Government initiative.

Data availability

Data will be made available on request.

References

[1] World Health Organization. World health statistics 2019: Monitoring health for the SDGs, sustainable development goals. Geneva: World Health Organization; 2019.
[2] World Health Organization. The power of cities: Tackling noncommunicable diseases and road traffic injuries. Technical report, Geneva: World Health Organization; 2019.
[3] Devailly F-X, Larocque D, Charlin L. Ig-Rl: Inductive graph reinforcement learning for massive-scale traffic signal control. IEEE Trans Intell Transp Syst 2021;23(7):7496–507.
[4] Li Z, Yu H, Zhang G, Dong S, Xu C-Z. Network-wide traffic signal control optimization using a multi-agent deep reinforcement learning. Transp Res Part C: Emerg Technol 2021;125:103059.
[5] Wang T, Cao J, Hussain A. Adaptive traffic signal control for large-scale scenario with cooperative group-based multi-agent reinforcement learning. Transp Res Part C: Emerg Technol 2021;125:103046.
[6] Yazdani M, Sarvi M, Asadi Bagloee S, Nassir N, Price J, Parineh H. Intelligent vehicle pedestrian light (IVPL): A deep reinforcement learning approach for traffic signal control. Transp Res Part C: Emerg Technol 2023;149:103991. http://dx.doi.org/10.1016/j.trc.2022.103991.
[7] Khamis M, Gomaa W, El-Shishiny H. Multi-objective traffic light control system based on Bayesian probability interpretation. In: 2012 15th international IEEE conference on intelligent transportation systems, intelligent transportation systems. ITSC, 2012 15th international IEEE conference on, IEEE; 2012, p. 995–1000. http://dx.doi.org/10.1109/ITSC.2012.6338853.
[8] Stevanovic A, Stevanovic J, So J, Ostojic M. Multi-criteria optimization of traffic signals: Mobility, safety, and environment. In: Engineering and applied sciences optimization (OPT-i) - professor Matthew G. Karlaftis Memoria. vol. 55, 2015, p. 46–68. http://dx.doi.org/10.1016/j.trc.2015.03.013, 55.
[9] Du W, Ye J, Gu J, Li J, Wei H, Wang G. Safelight: A reinforcement learning method toward collision-free traffic signal control. In: Proceedings of the AAAI conference on artificial intelligence. Vol. 37, 2022, p. 14801–10.
[10] Li X, Sun J-Q. Intersection multi-objective optimization on signal setting and lane assignment. Phys A 2019;525:1233–46. http://dx.doi.org/10.1016/j.physa.2019.04.223.
[11] Khattak MW, Pirdavani A, De Winne P, Brijs T, De Backer H. Estimation of safety performance functions for urban intersections using various functional forms of the negative binomial regression model and a generalized Poisson regression model. Accid Anal Prev 2021;151:105964. http://dx.doi.org/10.1016/j.aap.2020.105964.
[12] Yang D, Xie K, Ozbay K, Yang H. Fusing crash data and surrogate safety measures for safety assessment: Development of a structural equation model with conditional autoregressive spatial effect and random parameters. Accid Anal Prev 2021;152:105971. http://dx.doi.org/10.1016/j.aap.2021.105971.
[13] Wood JS, Donnell ET, Fariss CJ. A method to account for and estimate underreporting in crash frequency research. Accid Anal Prev 2016;95:57–66. http://dx.doi.org/10.1016/j.aap.2016.06.013.

[14] Tarko AP, Lizarazo CG. Validity of failure-caused traffic conflicts as surrogates of rear-end collisions in naturalistic driving studies. Accid Anal Prev 2021;149:105863. http://dx.doi.org/10.1016/j.aap.2020.105863.
[15] Yang D, Ozbay K, Xie K, Yang H, Zuo F. A functional approach for characterizing safety risk of signalized intersections at the movement level: An exploratory analysis. Accid Anal Prev 2021;163:106446. http://dx.doi.org/10.1016/j.aap.2021.106446.
[16] Ghoul T, Sayed T. Real-time signal-vehicle coupled control: An application of connected vehicle data to improve intersection safety. Accid Anal Prev 2021;162:106389. http://dx.doi.org/10.1016/j.aap.2021.106389.
[17] Wang K, Zhao S, Jackson E. Investigating exposure measures and functional forms in urban and suburban intersection safety performance functions using generalized negative binomial - P model. Accid Anal Prev 2020;148:105838. http://dx.doi.org/10.1016/j.aap.2020.105838.
[18] Kidando E, Kitali AE, Kutela B, Ghorbanzadeh M, Karaer A, Koloushani M, Moses R, Ozguven EE, Sando T. Prediction of vehicle occupants injury at signalized intersections using real-time traffic and signal data. Accid Anal Prev 2021;149:105869. http://dx.doi.org/10.1016/j.aap.2020.105869.
[19] Hu J, Huang M-C, Yu X. Efficient mapping of crash risk at intersections with connected vehicle data and deep learning models. Accid Anal Prev 2020;144:105665. http://dx.doi.org/10.1016/j.aap.2020.105665.
[20] Zheng L, Sayed T, Mannering F. Modeling traffic conflicts for use in road safety analysis: A review of analytic methods and future directions. Anal Methods Accid Res 2021;29:100142. http://dx.doi.org/10.1016/j.amar.2020.100142.
[21] Charly A, Mathew TV. Estimation of traffic conflicts using precise lateral position and width of vehicles for safety assessment. Accid Anal Prev 2019;132:105264. http://dx.doi.org/10.1016/j.aap.2019.105264.
[22] Peesapati LN, Hunter MP, Rodgers MO. Evaluation of postencroachment time as surrogate for opposing left-turn crashes. Transp Res Rec 2013;2386(1):42–51.
[23] Zheng L, Sayed T. A bivariate Bayesian hierarchical extreme value model for traffic conflict-based crash estimation. Anal Methods Accid Res 2020;25. http://dx.doi.org/10.1016/j.amar.2020.100111.
[24] Fu C, Sayed T. Identification of adequate sample size for conflict-based crash risk evaluation: An investigation using Bayesian Hierarchical extreme value theory models. Anal Methods Accid Res 2023;39:100281.
[25] Guo Y, Sayed T, Essa M. Real-time conflict-based Bayesian Tobit models for safety evaluation of signalized intersections. Accid Anal Prev 2020;144:105660. http://dx.doi.org/10.1016/j.aap.2020.105660.
[26] Hu Y, Li Y, Huang H, Lee J, Yuan C, Zou G. A high-resolution trajectory data driven method for real-time evaluation of traffic safety. Accid Anal Prev 2022;165. http://dx.doi.org/10.1016/j.aap.2021.106503.
[27] Hussain F, Ali Y, Li Y, Haque MM. Real-time crash risk forecasting using Artificial-Intelligence based video analytics: A unified framework of generalised extreme value theory and autoregressive integrated moving average model. Anal Methods Accid Res 2023;100302.
[28] Thorpe TL, Anderson C. Traffic light control using sarsa with three state representations. 1996.
[29] Abdulhai B, Pringle R, Karakoulas GJ. Reinforcement learning for true adaptive traffic signal control. J Transp Eng 2003;129(3):278–85.
[30] de Oliveira D, Bazzan AL, da Silva BC, Basso EW, Nunes L, Rossetti R, de Oliveira E, da Silva R, Lamb L. Reinforcement learning based control of traffic lights in non-stationary environments: A case study in a microscopic simulator. In: EUMAS. 2006.
[31] Kuyer L, Whiteson S, Bakker B, Vlassis N. Multiagent reinforcement learning for urban traffic control using coordination graphs. In: Machine learning and knowledge discovery in databases: European conference, ECML PKDD 2008, antwerp, Belgium, September 15-19, 2008, proceedings, Part I. vol. 19, Springer; 2008, p. 656–71.
[32] Wen K, Qu S, Zhang Y. A stochastic adaptive control model for isolated intersections. In: 2007 IEEE international conference on robotics and biomimetics. ROBIO, IEEE; 2007, p. 2256–60.
[33] Wiering MA. Multi-agent reinforcement learning for traffic light control. In: Machine learning: proceedings of the seventeenth international conference. ICML'2000, 2000, p. 1151–8.
[34] Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G. Human-level control through deep reinforcement learning. Nat 2015;518(7540):529–33.
[35] Genders W, Razavi S. Using a deep reinforcement learning agent for traffic signal control. 2016, arXiv preprint arXiv:1611.01142.
[36] Garg D, Chli M, Vogiatzis G. Deep reinforcement learning for autonomous traffic light control. In: 2018 3rd IEEE international conference on intelligent transportation engineering. ICITE, IEEE; 2018, p. 214–8.
[37] Sajad Mousavi S, Schukat M, Howley E. Traffic light control using deep policy-gradient and value-function based reinforcement learning. 2017, arXiv e-Prints arXiv:1704.08883.
[38] Oroojlooy A, Nazari M, Hajinezhad D, Silva J. Attendlight: Universal attention-based reinforcement learning model for traffic signal control. Adv Neural Inf Process Syst 2020;33:4079–90.
[39] Zhang S, Abdel-Aty M, Cai Q, Li P, Ugan J. Prediction of pedestrian-vehicle conflicts at signalized intersections based on long short-term memory neural network. Accid Anal Prev 2020;148:105799. http://dx.doi.org/10.1016/j.aap.2020.105799.
[40] Gokhale S, Pandian S. A semi-empirical box modeling approach for predicting the carbon monoxide concentrations at an urban traffic intersection. Atmos Environ 2007;41(36):7940–50. http://dx.doi.org/10.1016/j.atmosenv.2007.06.065.
[41] Li X, Sun J-Q. Effect of interactions between vehicles and pedestrians on fuel consumption and emissions. Phys A 2014;416:661–75. http://dx.doi.org/10.1016/j.physa.2014.09.028.
[42] Pandian S, Gokhale S, Ghoshal AK. Evaluating effects of traffic and vehicle characteristics on vehicular emissions near traffic intersections. Transp Res Part D: Transp Environ 2009;14(3):180–96. http://dx.doi.org/10.1016/j.trd.2008.12.001.
[43] Liang X, Du X, Wang G, Han Z. A deep reinforcement learning network for traffic light cycle control. IEEE Trans Veh Technol 2019;68(2):1243–53.
[44] Huang F, Liu P, Yu H, Wang W. Identifying if VISSIM simulation model and SSAM provide reasonable estimates for field measured traffic conflicts at signalized intersections. Accid Anal Prev 2013;50:1014–24.
[45] Fan R, Yu H, Liu P, Wang W. Using VISSIM simulation model and Surrogate Safety Assessment Model for estimating field measured traffic conflicts at Freeway Merge Areas.. IET Intell Transp Syst (Inst Eng Technol) 2013;7(1):68–77.
[46] Sobhani A, Young W, Sarvi M. A simulation based approach to assess the safety performance of road locations. Transp Res Part C: Emerg Technol 2013;32:144–58. http://dx.doi.org/10.1016/j.trc.2012.10.001.
[47] Guo Y, Sayed T, Zheng L, Essa M. An extreme value theory based approach for calibration of microsimulation models for safety analysis. Simul Model Pr Theory 2021;106:102172. http://dx.doi.org/10.1016/j.simpat.2020.102172.
[48] Guo Y, Essa M, Sayed T, Haque MM, Washington S. A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in Australia. Transp Res Part 2019;101:96–110.
[49] Katrakazas C, Quddus M, Chen W-H. A simulation study of predicting real-time conflict-prone traffic conditions. IEEE Trans Intell Transp Syst 2017;19(10):3196–207.
[50] Shahdah U, Saccomanno F, Persaud B. Application of traffic microsimulation for evaluating safety performance of urban signalized intersections. Transp Res Part 2015;60:96–104.
[51] Federal Highway Administration (FHWA). Surrogate Safety Assessment Model. Federal Highway Administration (FHWA); 2017.
[52] Hayward JC. Near miss determination through use of a scale of danger. University Park, Pa.: Pennsylvania Transportation and Traffic Safety Center, The Pennsylvania State University; 1972.

[53] Cooper PJ. Experience with traffic conflicts in Canada with emphasis on ‘‘post encroachment time’’ techniques. In: Asmussen E, editor. International calibration study of traffic conflict techniques. Berlin, Heidelberg: Springer Berlin Heidelberg; 1984, p. 75–96. http://dx.doi.org/10.1007/978-3-642-82109-7_8.
[54] Mahmud SMS, Ferreira L, Hoque MS, Tavassoli A. Micro-simulation modelling for traffic safety: A review and potential application to heterogeneous traffic environment. IATSS Res 2019;43(1):27–36. http://dx.doi.org/10.1016/j.iatssr.2018.07.002.
[55] Johnsson C, Laureshyn A, Dágostino C. A relative approach to the validation of surrogate measures of safety. Accid Anal Prev 2021;161:106350. http://dx.doi.org/10.1016/j.aap.2021.106350.
[56] Bahouth G, Graygo J, Digges K, Schulman C, Baur P. The benefits and tradeoffs for varied high-severity injury risk thresholds for advanced automatic crash notification systems. In: Traffic injury prevention. vol. 15, Great Britain: Taylor & Francis; 2014, p. S134–40.
[57] Brumbelow M. Front crash injury risks for restrained drivers in good-rated vehicles by age, impact configuration, and EDR-based delta V. In: Conference proceedings international research council on the biomechanics of injury. IRCOBI, International Research Council on the Biomechanics of Injury; 2019, p. 561–75.
[58] Joksch HC. Velocity change and fatality risk in a Crash—A rule of thumb. Accid Anal Prev 1993;25(1):103–4. http://dx.doi.org/10.1016/0001-4575(93)90102-3.
[59] Laureshyn A, De Ceunynck T, Karlsson C, Svensson Å, Daniels S. In search of the severity dimension of traffic events: Extended Delta-V as a traffic conflict indicator. Accid Anal Prev 2017;98:46–56. http://dx.doi.org/10.1016/j.aap.2016.09.026.
[60] Shannon D, Murphy F, Mullins M, Rizzi L. Exploring the role of delta-V in influencing occupant injury severities – A mediation analysis approach to motor vehicle collisions. Accid Anal Prev 2020;142. http://dx.doi.org/10.1016/j.aap.2020.105577.
[61] World Health Organization. Global status report on road safety 2018. Geneva: World Health Organization; 2018.
[62] Altendorfer R, Wilkmann C. A new approach to estimate the collision probability for automotive applications. Autom 2021;127. http://dx.doi.org/10.1016/j.automatica.2021.109497.
[63] Hussain F, Li Y, Arun A, Haque MM. A hybrid modelling framework of machine learning and extreme value theory for crash risk estimation using traffic conflicts. Anal Methods Accid Res 2022;36. http://dx.doi.org/10.1016/j.amar.2022.100248.
[64] Oh C, Kim T. Estimation of rear-end crash potential using vehicle trajectory data. Accid Anal Prev 2010;42(6):1888–93. http://dx.doi.org/10.1016/j.aap.2010.05.009.
[65] Tarko AP. A unifying view on traffic conflicts and their connection with crashes. Accid Anal Prev 2021;158:106187. http://dx.doi.org/10.1016/j.aap.2021.106187.
[66] Hensher DA. Value of Life and Injuries. In: Vickerman R, editor. International encyclopedia of transportation. Oxford: Elsevier; 2021, p. 737–41. http://dx.doi.org/10.1016/B978-0-08-102671-7.10205-2.
[67] Hensher DA, Rose JM, Ortúzar Jd, Rizzi LI. Estimating the willingness to pay and value of risk reduction for car occupants in the road environment. Transp Res Part A 2009;43(7):692–707. http://dx.doi.org/10.1016/j.tra.2009.06.001.
[68] Jou R-C, Hensher DA, Chen T-Y, Chao M-C. Hospitalisation costs and duration of elderly motorcyclists’ non-fatality crashes in Taiwan. Int J Inj Control Saf Promot 2013;20(2):158–68. http://dx.doi.org/10.1080/17457300.2012.720579.
[69] Transport for New South Wales. Transport for NSW economic parameter values. Technical report, Sydney, NSW: TfNSW; 2020.
[70] Huang H, Zeng Q, Pei X, Wong SC, Xu P. Predicting crash frequency using an optimised radial basis function neural network model. In: Transportmetrica.a.transport science. vol. 12, Great Britain: Taylor & Francis; 2016, p. 330–45.
[71] Mafi S, Abdelrazig Y, Doczy R. Analysis of PGap Acceptance Behavior for unprotected right and left turning maneuvers at signalized intersections using data mining methods: A driving simulation approach. Transp Res Rec 2018;2672(38):160–70. http://dx.doi.org/10.1177/0361198118783111.
[72] Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM sigkdd international conference on knowledge discovery and data mining. 2016, p. 785–94. http://dx.doi.org/10.1145/2939672.2939785, arXiv:1603.02754.
[73] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: Machine Learning in Python. J Mach Learn Res 2011;12:2825–30.
[74] Yazdani M. Multi-modal traffic signal control using deep reinforcement learning [Ph.D. thesis], The University of Melbourne; 2023.