Cycling safety assessment in microscopic traffic simulation: A review and methodological framework

Konrad, Christoph M. ; Dabiri, A.; Schulte, F.; Moore, J.K.; Happee, R.

DOI
10.1016/j.trip.2025.101734

Publication date
2025

Document Version
Final published version

Published in
Transportation Research Interdisciplinary Perspectives

Transportation Research Interdisciplinary Perspectives 34 (2025) 101734

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

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

Cycling safety assessment in microscopic traffic simulation: A review and methodological framework

Christoph M. Konrad *, Azita Dabiri, Frederik Schulte, Jason K. Moore, Riender Happee

Faculty of Mechanical Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

ARTICLE INFO

Keywords:
Cycling safety
Microscopic traffic simulation
Surrogate safety assessment
Causality
Validation
Vulnerable road users

ABSTRACT

Vulnerable road user safety is paramount for increasing shares of active travel modes and introducing automated vehicles. Microscopic traffic simulation is a prevalent method in research and practice with a growing focus on safety and cyclists. Its practical benefits make it an essential tool for developing safe future transportation. We review the methodology of simulation studies and the validation of their microscopic models to evaluate cycling safety assessment in microscopic simulations. We find that current work relies predominantly on the lane-based models of established traffic flow simulation packages that separate longitudinal and lateral dynamics. These models do not sufficiently capture diverse behaviors and conflict causality to predict cycling safety. In contrast, new models with successful calibrations and validations advance simulated interactions towards capturing conflict causality. Of 42 reviewed studies, six calibrate, and three validate models for safety prediction. Other studies disregard calibration and validation, posing a threat of unfounded safety predictions and unsafe design recommendations. We present a methodological framework conceptualizing best practices for reliable assessment. It calls for the identification of safety-relevant behaviors of cyclists and other road users in conflicts. Specialized behavioral models must be developed, calibrated, and validated. The selected safety indicators must enable capturing the expected unsafe events. To create these tools, improved models of cycling behavior must be transferred to established simulation packages. Following the framework, researchers and practitioners can use simulation as a practical and ethical means to assess the cycling safety impact of innovations ranging from infrastructure to automation and connectivity.

1. Introduction

Urban cycling is integral to sustainable transportation futures like the E-bike City (Ballo et al., 2023). Globally, bicycle traffic in urban areas is increasing (Pucher and Buehler, 2021). Crucially, though, cyclists carry a high burden regarding traffic safety. In the Netherlands in 2021, cyclists sustained 71% of all serious traffic injuries (Aarts et al., 2022). Bicycle-oriented traffic systems and infrastructure designs can improve cycling safety and thus are central to recent transport policies. At the same time, Automated Vehicles (AVs) promise increased traffic safety but are especially challenged by interactions with cyclists and other vulnerable road users.

Microscopic traffic simulation is a prevalent tool in traffic engineering and research. It recreates time-varying traffic phenomena by simulating individual agents in a virtual environment. Waiting for accidents to occur in real traffic takes long observation periods and can be considered unethical (Essa and Sayed, 2015; Laureshyn et al., 2017). In contrast, the ability to test innovations in accelerated virtual environments, without physical implementation, and without exposing subjects to harm, makes simulation an ethical and practical tool for early development stages. Researchers use it to assess infrastructure, traffic control, and emerging technologies like AVs or intelligent transport systems. Additionally, it serves as a training environment for learning- and optimization-based algorithms. Especially AV research relies on integrated simulation environments to develop vehicle-level algorithms or logistics applications and predict their safety impact. However, simulation results inherit the uncertainties of underlying model assumptions (Sohrabi et al., 2021). An accurate representation of cyclist behavior is required to ensure that the innovations developed, trained, and tested with micro-simulation are robust, reliable, and guarantee bicycle-friendly traffic design.

Microscopic simulations are designed to model operational characteristics like vehicle flow, delay, trip duration, and congestion. For simplification, they are mostly collision-free and omit crash causality. Still, the simulated trajectories technically enable the calculation of

* Corresponding author.
E-mail addresses: c.m.konrad@tudelft.nl (C.M. Konrad), a.dabiri@tudelft.nl (A. Dabiri), f.schulte@tudelft.nl (F. Schulte), j.k.moore@tudelft.nl (J.K. Moore), r.happee@tudelft.nl (R. Happee).

https://doi.org/10.1016/j.trip.2025.101734
Received 20 February 2024; Received in revised form 22 October 2025; Accepted 9 November 2025
Available online 16 November 2025

surrogate safety measures based on the nearness of Road Users (RUs) in conflicts, regardless of simulated collisions. Arun et al. (2021) review crash surrogacy and report that a definitive validation of the crash-conflict relationship is missing, even for natural traffic observations. For simulation, they point out the insufficiency of existing behavioral models to describe unsafe interactions. While Young et al. (2014) also observe model limitations, they highlight the successful correlation of simulated and observed conflicts in specific automobile traffic scenarios achieved by various calibration studies (Huang et al., 2013; Gettman et al., 2008; Guo et al., 2019). Despite the good overall fit, though, Guo et al. (2019) also observe a poor correlation of simulated and observed conflict locations. Tarko (2018) suggests that exposure alone may cause the good overall fit.

Cycling behavior differs from the disciplined lane-based behavior of cars. Cyclists exploit the flexibility of their mobile vehicles to use infrastructure in various, sometimes unintended ways. For example, cyclists cross intersections with direct and indirect left turns, ride against traffic, and use car or pedestrian facilities (Twaddle and Busch, 2019). Discontinuities in dedicated cycling infrastructure lead to riding outside the intended infrastructure (Nabavi Niaki et al., 2018). Bicycle boxes, two-directional lanes, or protected intersections may require maneuvers that do not have an equivalent in car traffic. Cyclists exhibit riding in groups (Grigoropoulos et al., 2021) and queuing patterns at traffic lights (Gavriilidou et al., 2019a) that distinguish them from cars. Interactions with AVs create new cycling behaviors ranging from caution and consideration to exploitation (Bjørnskau et al., 2023). Despite these complex and unique behaviors, many current simulation environments reuse car models with different parameter values to represent cycling. Researchers have raised concerns if these models can create valid assessments of real-world situations (Twaddle et al., 2014). This is especially critical for safety, which, after Tarko (2018), is determined by individual behaviors and interactions rather than aggregated measures. Johnsson et al. (2018) review surrogate safety assessment for vulnerable RUs and conclude that the suitability of indicators depends on their ability to capture the safety-relevant interactions. With rising interest in cycling, simulation-based safety research articles, case studies, and government-issued reports are increasingly being published. To analyze cycling safety on varying infrastructure (intersections, roads, dedicated paths, Intelligent Transport System (ITS) scenarios) and with varying interaction partners (cyclists, pedestrians, conventional and automated cars), the corresponding models must describe complex behaviors. However, legitimate concerns exist about the capability of micro-simulation to accurately model cycling behavior and safety. This literature review analyzes the methods of existing simulation-based cycling safety studies. We compare these with best practices from the micro-simulation domain and highlight differences regarding cycling. We track where researchers observe shortcomings that limit their simulation results. To this end, we address the subsequent research questions:

RQ1: Can micro-simulation model the safety-relevant behavior and riding dynamics of cyclists?

RQ2: Can micro-simulation models be calibrated and validated to predict traffic conflicts involving cyclists?

RQ3: Can cycling safety be evaluated based on micro-simulation?

Assuming that actual traffic safety is a product of RU behaviors and interactions, we review the elements relevant for creating, executing, and assessing simulated interactions. Based on the three dimensions of human–vehicle interaction simulations (Janssen et al., 2020), these are the simulation software package defining the environment, the cyclist models defining the agents' behaviors, and the scenarios defining the interactions. We add the calibration and validation procedures that ensure reliability and the safety quantification techniques.

Based on the review results, we present a framework that thoroughly discusses requirements for reliable study results on top of existing guidelines for simulation-based operational traffic assessment and highlights research gaps. Our contributions are two-fold. Firstly, our framework provides researchers and practitioners with high-level methodological best practices to avoid unreliable results. Secondly, we identify the current limitations and highlight research needs to elevate micro-simulation to a tool that fully integrates bicycle traffic and adequately captures safety aspects.

Section 2 describes the review method. Section 3 overviews the included studies and analyzes their simulated cycling safety assessment methods. Section 4 derives a methodological framework, discusses requirements for a robust assessment, and proposes a research agenda. Finally, Section 5 provides concluding remarks and an outlook on future work.

2. Literature review method

We conduct a systematic literature review of micro-simulation studies that assess cycling safety and analyze the methods of these publications regarding our predefined selection of aspects that govern simulated interactions. If necessary, we complement results with insights from the general domain (Fig. 1).

The subsequent criteria delimit the review scope.

Microscopic Traffic Simulation: Design and application of traffic simulation by modeling individual RUs. We only include studies with predictive behavioral models. This focuses on fully virtual agents with high fidelity in the framework of Janssen et al. (2020). It excludes human-in-the-loop simulations and static trajectory models based on field or driving simulator data.

Quantification of Safety: Assessment of actual traffic safety related to incident or injury risk using quantifiable criteria. We do not consider perceived safety.

Cycling: Inclusion of two-wheelers that require pedaling as propulsion. While we include pedal support (e-bikes), we exclude motorized two-wheelers and motorcycles that do not require pedaling due to their different dynamic characteristics. If a clear definition is absent, we rely on basic terminology like "cyclist" or "bicycle" in the papers.

We search titles, abstracts, and keywords in Scopus, IEEE Xplore, Web of Science, and TRID. Our search phrases consist of synonyms of the inclusion criteria microscopic traffic simulation (micro-simulation, microscopic simulation, traffic simulation, agent-based simulation, simulating traffic), cycling (bicycle, bike, bicyclist, cyclist, pedelec, tricycle, trike), and quantification of safety (safety, conflict, crash, accident, collision, injury). We combine synonyms with Boolean AND, aspects with Boolean OR, and use wildcard and proximity operators. Due to syntax requirements, the actual search queries differ between databases.

We limit the search results to journals, conference proceedings, and research/technical reports and eliminate duplicates. Then, we filter by title, abstract, and full text. We search the internet, use library services, and contact the authors to obtain full-text access. If necessary, we machine-translate articles into English. We exclude articles as soon as it becomes clear that the inclusion criteria above are not satisfied. Finally, we add articles previously known to the authors and perform snowballing with Google Scholar for all included papers. Our search cut-off is December 08, 2023.

3. Literature review results

This section presents an overview of the included studies (Section 3.1) and the individual results for each dimension of simulated cycling safety assessment.

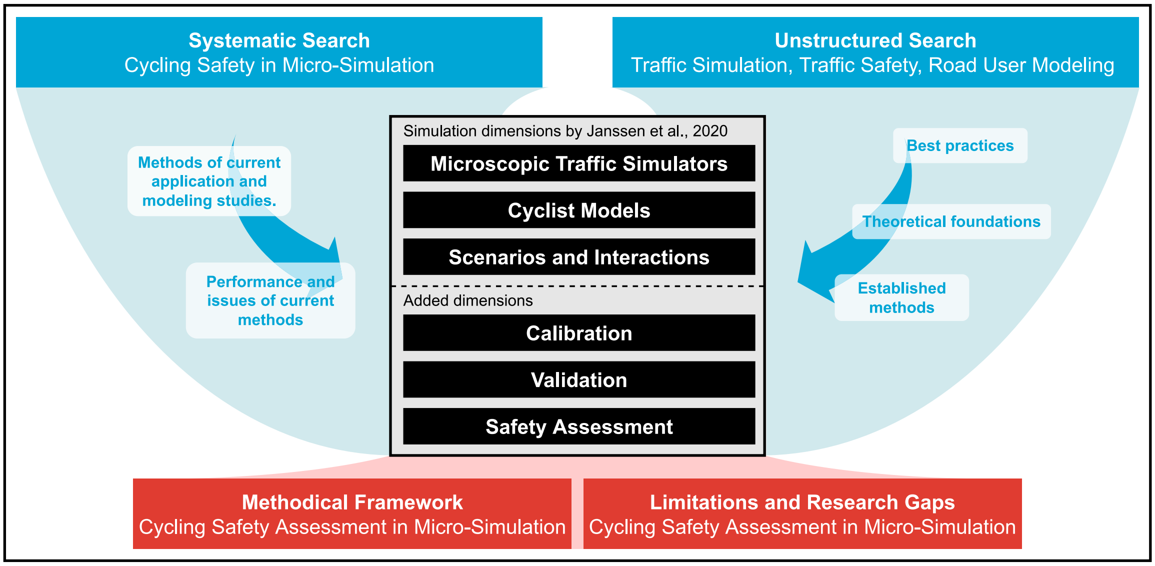
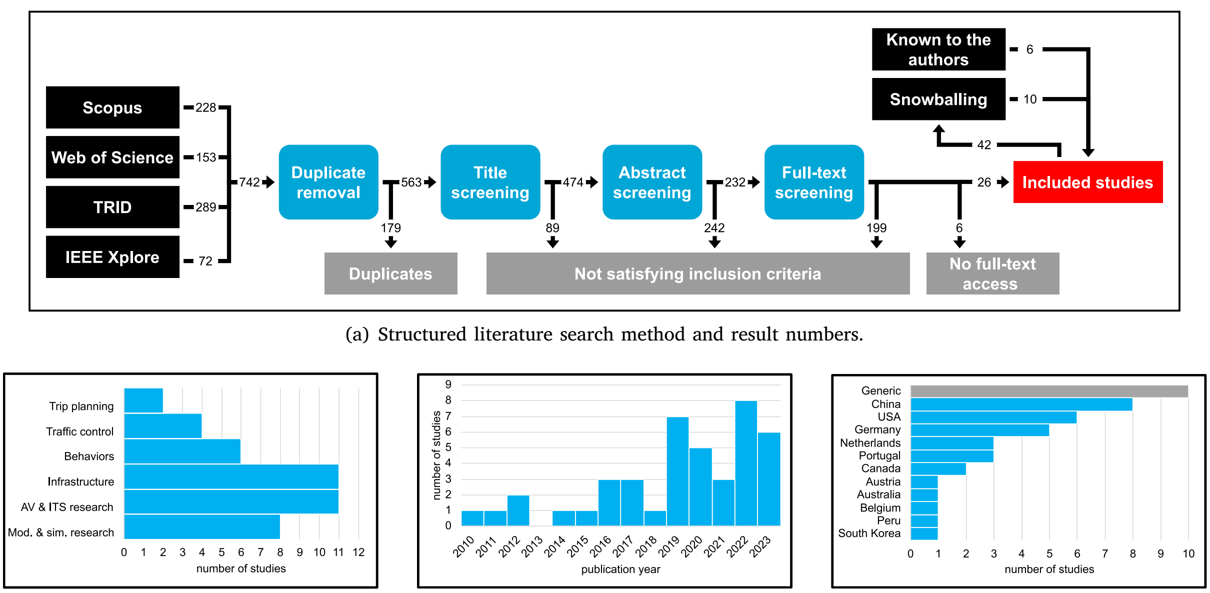


Fig. 1. Overview of the review method with systematic and unstructured search (blue, top) generating the results (black, center) categorized by the dimensions of simulated cycling safety assessment. Discussion of the results leads to the outcome (red, bottom).



(a) Structured literature search method and result numbers.

(b) Research question topic categories.

(c) Publication years.

(d) Study locations.

Fig. 2. Overview of the search results.

3.1. Overview of included studies

After screening all search results and performing snowballing, we include 42 studies (Fig. 2a). Table 1 summarizes the included studies. We identified six topic categories (Fig. 2b). "Mod. & sim. research" describes publications that focus on developing, calibrating, and validating simulation models. Among these, most develop new cyclist models outside or linked to existing simulation packages (hereafter called "new external models"). All other studies apply existing models to answer safety-related research questions (hereafter called "application studies") and are categorized based on their focus on traffic control, trip planning, infrastructure elements, AV & ITS research, and specific behaviors.

Publication numbers increased in recent years, highlighting the relevance of the topic (Fig. 2c). As the traffic characteristics vary globally, Fig. 2d lists the study location. While eleven countries allow us to observe some regional context, the results do not provide a complete global picture. Ten studies simulate a generic scenario without a corresponding real-world location and data-driven calibration or validation. Only six calibrate, and only three validate their models in terms of safety.

3.2. Microscopic traffic simulation packages

The simulation environment is defined by the fundamental implementation of its traffic network elements. Over half of the reviewed studies use the software packages SUMO (8 of 42) and VISSIM (20 of 42). Both are traffic flow simulators focusing on the accurate representation of operational characteristics. They provide the fundamental functionalities to include cyclists, albeit with limitations in their models. The next biggest category (6 of 42) comprises studies with custom

Table 1
Reviewed studies using micro-simulation to assess cycling safety through exposure, simulated collisions, or surrogate safety assessment (SSA). The interaction opponents considered for safety analysis are bicyclists (B), e-bicyclists (EB), human-driven heavy vehicles (HV), human-driven cars (HC), automated cars (AC), pedestrians (P), and other (O). (*) indicates not clearly defined cases.

Reference	Category	Package	Interaction opponent							Scenario	Safety method	Predictive validation	Study objective
			B	EB	HV	HC	AC	P	O				
			Traffic Analysis Applications										
Denk et al. (2022)	AV & ITS	Custom				x	x			Right-hook conflict at protected bike lanes, Germany	collisions		Assessment of V2X functions to avoid conflicts in right-hook scenarios.
Jiang et al. (2022)	AV & ITS	SUMO					x			Major intersection with bicycle lanes	SSA		Train AVs driving policies based on simulated traffic evaluating a distance-based safety cost.
Karkhanis et al. (2020)	AV & ITS	SUMO			x					Unsignalized minor intersection, Netherlands	collisions		Analyze an ITS warning system at a bicycle path and bus road crossing.
Pauwels et al. (2022)	AV & ITS	SUMO				x	x			Large urban area, Netherlands	SSA		Analyze different AV driving characteristics and connectivity levels for VRU safety.
Pechinger et al. (2021)	AV & ITS	AIMSUN					x			Intersection with protected bike lanes, Germany	collisions		Analyze conflicts and collisions created by a Hardware-in-the-loop simulation of AV systems.
Pechinger et al. (2023)	AV & ITS	AIMSUN					x			Intersection with protected bike lanes and parking, Germany	collisions		Analyze conflicts and collisions created by a Hardware-in-the-loop simulation of AV systems and infrastructure based perception.
Qian et al. (2022)	AV & ITS	SUMO			x	x				Crossing and right-hook conflicts	SSA		Analyze various C-ITS applications for cyclist and pedestrian safety.
Ren et al. (2023)	AV & ITS	SUMO					x			Major intersection with bike lanes	collisions		Train and assess adversarial policies for AV control on simulation output in terms of safety.
Ren et al. (2022)	AV & ITS	SUMO					x			Major intersection with bike lanes	SSA		Train and assess environment encoding for AV control on simulation output in terms of safety.
Tafidis et al. (2019)	AV & ITS	VISSIM			x		x			Neighborhood with shared roads, Belgium	SSA		Assess AV driving characteristics for cycling safety.
Thompson et al. (2020)	AV & ITS	Netlogo				x	x			Abstract grid-based road network	collisions		Simulate the safety effect of behavioral adaptation to perfect AVs.
Bahmankhah et al. (2019b)	Behaviors	VISSIM				x				Multi-lane roundabouts, Portugal	SSA		Analyze impact of driving volatility.
Li et al. (2011)	Behaviors	Custom				x				Minor road with bike lanes	SSA		Analyze impact of cyclists using vehicle lanes to overtake.
Ren et al. (2016)	Behaviors	Custom	x	x						Through-going bike lanes at major intersections, China	SSA		Analyze cyclist dispersion effect while crossing intersections.
Thompson et al. (2015)	Behaviors	Netlogo				x				Abstract grid-based road network	exposure		Simulate behavioral adaptation to exposure to explain the safety-in-numbers effect.
Thompson et al. (2016)	Behaviors	Netlogo				x				Abstract grid-based road network	collisions		Simulate behavioral adaptation to local density to explain the safety-in-numbers effect.
Wallentin and Loidl (2016)	Behaviors	Netlogo	x							Large urban area, Austria	exposure	operational	Analyze how exposure explains accident statistics.
Bahmankhah et al. (2019a)	Infrastructure	VISSIM				x				Major intersection and roundabout, Portugal	SSA		Compare safety of different roundabout layouts.
Campisi et al. (2020)	Infrastructure	VISSIM			x	x				Turbo roundabouts	SSA		Analyze the relationship between operational performance and safety of turbo roundabouts.
Dijkstra (2012)	Infrastructure	S-PARAMICS				x				Bike path crossing, Netherlands	SSA		Analyze impact of traffic redirection.
Guhathakurta et al. (2023)	Infrastructure	VISSIM	x			x*				Large urban area, USA	SSA		Analyze a cycling network design algorithm.
Joo et al. (2012)	Infrastructure	VISSIM	x		x*	x				Major intersection with bicycle left-turn lanes, bicycle boxes and shared roads, South Korea	SSA		Analyze performance of different left-turn cycling infrastructures.

(*continued on next page*)

implementations. This section summarizes the two most popular environments and briefly reviews their cyclist and safety assessment support.

Most researchers use the commercial software PTV VISSIM (PTV Group, 2023). Fellendorf and Vortisch (2011) describe its lane-based longitudinal and lateral continuous movement model. Along these

Table 1 (*continued*).

Kodupuganti and Pulugurtha (2022)	Infrastructure	VISSIM	x			x	x	x	Arterial road with bicycle lanes and light rail, USA	SSA		Analyze impact of a proposed light-rail system.
Monsere et al. (2019)	Infrastructure	VISSIM			x	x			Intersections with various shared turning facilities, USA	SSA		Analyze relationship between load and safety of existing cycling infrastructure.
Preston and Pulugurtha (2021)	Infrastructure	VISSIM	x		x	x			Protected intersection design, USA	SSA		Predict safety gains of protected intersection designs.
Russo et al. (2022)	Infrastructure	VISSIM				x			Intersections with various shared and mixed turning facilities, USA	SSA		Assess different shared and separated cyclist crossing facilities under varying RU volumes.
Silva et al. (2023)	Infrastructure	VISSIM	x		x	x		x	Major intersection with shared and dedicated lanes, Peru	SSA		Assess alternative cyclist crossing facility and adaptive traffic control at existing intersection.
Thompson et al. (2017)	Infrastructure	Netlogo				x			Abstract grid-based network with dedicated roads	exposure		Simulate the safety effect of behavioral adaptation to separated cycling infrastructure.
Alecsandru et al. (2010)	Traffic control	VISSIM				x			Arterial road with bicycle lanes, Canada	exposure		Optimize signal control for cycling safety.
Ledezma-Navarro et al. (2018)	Traffic control	VISSIM			x	x			Variations of bike lanes at major intersections, Canada	SSA		Compare different control strategies.
Lu and He (2019)	Traffic control	VISSIM	x*			x*	x		Neighborhood with shared and dedicated infrastructure, China	SSA		Analyze proposed treatment to traffic control in a school environment.
Wu et al. (2014)	Traffic control	VISSIM				x			Major intersection with bike lanes, China	exposure		Analyze cyclist volume adaptive signal control.
Bahmankhah and Coelho (2017)	Trip planning	VISSIM				x			Neighborhood with shared roads, Portugal	SSA		Determine optimal cycling routes with multiple objectives.
Hübner et al. (2017)	Trip planning	SUMO							Large urban area, Germany	exposure		Introduce and test cycling routing application.
Modeling and Simulation Research												
Langer et al. (2023)	Mod. & sim.	SUMO				x			Large urban area, Germany	collisions		Introduce perception and crash severity models and calibrate to statistics from injury and crash databases.
Lemcke et al. (2021)	Mod. & sim.	VISSIM			x	x			Right-hook conflict at turning lanes, USA	SSA		Investigate model parameter sensitivity and performance for cycling safety.
Li et al. (2020)	Mod. & sim.	Custom	x	x		x*			Major road with bike lanes, China	SSA		Introduce and test cellular automata overtaking model.
Liu et al. (2020)	Mod. & sim.	TESS				x*			Major road with bike lanes, China	SSA	safety, behaviors, operational	Introduce and test lane-based overtaking model.
Ni et al. (2023)	Mod. & sim.	Custom	x	x		x			Major road with bike lanes, China	SSA	behaviors, operational	Introduce and test social-force interaction model with Bayesian decision network.
Sun et al. (2019)	Mod. & sim.	Custom, VISSIM, TransMod.	x		x	x		x	Minor road with shared lanes and bus stops, China	SSA		Introduce state-machine overtaking model and compare to VISSIM and TransModeler.
Sun et al. (2020)	Mod. & sim.	VISSIM				x*			Major intersection with bike lanes, China	SSA	safety	Introduce deep learning vehicle left-turn model, train on VISSIM simulations, and validate on field observations.
Xu et al. (2023)	Mod. & sim.	VISSIM			x	x			Major intersection and roundabout with shared facilities, Australia	SSA	safety, behaviors, operational	Introduce and compare VR human-in-the-loop simulation with pure VISSIM.

lanes, car-following models govern longitudinal motion. Laterally, vehicles optimize their continuous on-lane position for maximum Time-To-Collision (TTC). A lane-change model governs the decision to transition into an adjacent lane to achieve free driving. At intersections, RUs respect priority rules. Bicycles and cars use the same models with different parameter values, but bicycles have a diamond-shaped footprint to improve queuing. The COWI (2012) report identified VISSIM settings that best approximate field-measured cyclist capacity, travel time, and delay over various mixed-traffic and dedicated infrastructures. The derived COWI (2013) report provides the standard guideline for cycling simulation with VISSIM. The proposed settings are not validated for safety assessment. Still, VISSIM technically supports trajectory export for automated safety assessment with the Surrogate Safety Assessment Model (SSAM) by Gettman et al. (2008).

The open-source framework SUMO (Lopez et al., 2018) models continuous longitudinal and discrete lateral dynamics along fixed lanes. A sublane model for continuous lateral dynamics (Semrau et al., 2016) and multiple car-following models are available. At intersections, vehicles consider routes, traffic laws, and collision avoidance. SUMO's support for cycling simulation is under development. Currently, bicycles reuse car-following models with adapted parameter values. SUMO supports cyclist-specific infrastructure and priority rules. While the model is designed to be collision-free, behavioral parameters tuned to unsafe driving may provoke collisions (German Aerospace Center, 2023b). A workshop of SUMO developers, users, and researchers during the SUMO User Conference 2022 (Roosta et al., 2023) discussed SUMO's cyclist models and concluded that the included car-following models, particularly the Intelligent Driver model (Kesting et al., 2010), can sufficiently represent cyclist following behaviors for current (typically not safety-related) applications. However, it was observed that the models should be more rigorously calibrated and expanded to include cycling-specific behaviors like side-by-side riding. Further conclusions suggest that the lateral behavior and the treatment of conflict areas lack the variability of real cycling. Cyclists strictly adhere to (sub)lanes and do not enter conflict areas if occupied by other road users. Finally, SUMO developers highlighted during the workshop that they did not validate SUMO's current cyclist simulation framework against data. Still, virtual logging devices enable to record surrogate safety indicators (German Aerospace Center, 2023c).

3.3. Cyclist models

Cyclist models describe the dynamics and behaviors of simulated agents. The reviewed application studies generally rely on the models included in existing software packages. Frequently, they adapt model parameters to create the desired behaviors in specific conditions.

Longitudinal behavior

Both SUMO and VISSIM reuse lane-based car-following models like the Wiedemann-74 (W74) (Wiedemann, 1974), Wiedemann-99 (W99), Krauss (Krauß, 1998), or Intelligent Driver (Kesting et al., 2010) models for the longitudinal behavior of bicycles. Among the reviewed studies, Tafidis et al. (2019) select W99 over W74, referencing increased flexibility in the larger number of parameters. Based on field observations, they create two parameter sets for leading cars and leading bicycles. Their adjustments aim to recreate the distances, cautiousness on shared roads, and sensitive driving reactions from Belgian cities. Xu et al. (2023) calibrate W99 for cycling on intersections and roundabouts. In their integrated simulation environment with additional behavioral sub-models, they highlight the importance of car-following for the overall performance and observe similar distance-vs.-time trajectories in simulation and reality. Other researchers select W74 car-following to recreate cycling at intersections in the United States (Russo et al., 2022; Lemcke et al., 2021) or cycling on intersections and roundabouts in Portugal (Bahmankhah et al., 2019b,a). In a comparison of field-measured and simulated conflicts, Lemcke et al. (2021) determine several sensitive W74 parameters, with the average standstill distance having the most significant impact on simulated safety.

Additionally, researchers adjust speed and acceleration to influence longitudinal dynamics. VISSIM enables definitions of speed distributions per road section and per vehicle type. Studies define different distributions for road sections and intersections (Monsere et al., 2019) and road sections and turns (Tafidis et al., 2019). Only Bahmankhah et al. (2019b) set the maximum deceleration to field observations. Although Grigoropoulos et al. (2022) observe that the acceleration from a stop at an intersection directly determines how long cyclists are present in safety-critical conflict zones, no study based on existing models adapts the default acceleration profiles.

Lateral behavior

SUMO and VISSIM model lateral behavior separately from longitudinal behavior. Where movements between lanes are significant, like for cycling on two-lane roundabouts (Bahmankhah et al., 2019a,b), researchers tune lane-change parameters. For on-lane behavior in VISSIM, Tafidis et al. (2019) adapt the preferred position on the lane, diamond-shaped footprints, the minimum lateral distance, observation of adjacent lanes, and collision time gain. In a VISSIM calibration study, Kaths et al. (2021) suggest that models for lateral behavior are more important for achieving realistic cycling than longitudinal behavior.

Human variance and error

Microscopic models are typically stochastic to recreate the inherent variability of road traffic. To introduce the increased variability of human errors, some reviewed studies modify traffic-rule adherence, cyclist attention, and perception. For example, Langer et al. (2023) introduce a perception error model which describes probabilities of RUs' failures to recognize a conflict and give priority. They calibrate a large-scale SUMO simulation to reproduce real-world crash and injury statistics. Additionally, they configure varying reaction times. They limit evasive maneuvers to braking and maintain SUMO's lane-based dynamics. While the simulation achieves good scores for car-car crash distributions, higher residuals remain for car-bicycle crashes. See Section 3.6 for details on simulated collisions.

Within the parameter space offered by existing simulation packages, researchers configure road users to ignore a set of traffic rules to simulate reckless or inattentive riding (Xu et al., 2023; Karkhanis et al., 2020). Other VISSIM studies adapt the temporary lack of attention, preferred safety distances, gap acceptance, and visibility distances to create unsafe behaviors (Lemcke et al., 2021; Bahmankhah et al., 2019b).

Models outside existing software packages.

Next to the standard lane-based models in existing software packages, researchers use and develop new individual or external models.

Thompson et al. (2015, 2016, 2017, 2020) and Wallentin and Loidl (2016) reduce cycling dynamics to longitudinal movement with minimal local interactions. Instead, they focus on software agents with independent intelligence and accurate exposure modeling to analyze exposure-accident relationships. Similarly, Denk et al. (2022) simulate and overlay undisturbed trajectories to create initial conditions for stochastic models of the encounter outcome. A related approach outside the scope of our review is the simulation of cyclists following fixed field-measured trajectories (Ma et al., 2017; Ni et al., 2019). This guarantees accurate, undisturbed trajectories, but the simulated cyclists cannot react to their environment. While it enables the counterfactual analysis of vehicle functions (e.g., Zhou and Wang, 2022) when the reactions of the cyclists are not relevant, longer, complex RU interactions cannot be simulated.

In contrast, other researchers extend the lane-based dynamics of popular simulation packages. A fundamental assumption of lane-based models in existing simulation packages is that RUs follow a dominant direction of travel. This generally holds on road segments, but is inadequate with unordered traffic at intersections. There, cyclists show varying behaviors regarding left turns, the direction of travel, or the response to signals (Twaddle and Busch, 2019). For automobiles, researchers show that the lane-based architecture constrains the lateral movement diversity so that varying turning trajectories, specifically in interactions with cyclists, are insufficiently captured (Ni et al., 2019; Ma et al., 2017). Among the reviewed studies, Xu et al. (2023) create a VISSIM-Unity co-simulation to improve the lateral behavior of cyclists in VISSIM with Unity's velocity-obstacle path-finding algorithm. A 3D environment also enables them to simulate weather effects that change the vehicle dynamics. Next, studies use social force models to achieve spatially fully continuous two-dimensional interactions (Li et al., 2011; Ni et al., 2023). This model class describes cyclists as particles moved by imaginary forces exerted by the environment, intentions, and other road users. In our previous work (Schmidt et al., 2024), a two-wheeler dynamics model based on social forces is proposed that can respect the motion constraints imposed by a bicycle and rider. It uniquely introduces explicit steer, yaw, and roll dynamics for micro-simulated cyclists, making only physically feasible maneuvers possible. The model shows promising performance for cyclist interaction examples but has not yet been validated. Finally, the cellular automata used by Ren et al. (2016) and Li et al. (2020) also enable extended movements in the lateral dimension. Other than the social force approach, they are spatially discrete.

Another issue is the prediction of consistent maneuver decisions over varying time horizons. Here, reviewed studies propose Bayesian models (Liu et al., 2020; Ni et al., 2023), finite state machines (Sun et al., 2020), and other rule-based approaches (Ren et al., 2016; Li et al., 2020). For car behavior in conflicts with cyclists, Sun et al. (2020) create a deep-learning-based path planning model. Outside our reviewed studies, researchers overlay social-force-based models with rule-based models (Rinke et al., 2017) to describe perception and decision-making. Kaths (2023) model the tactical choice between direct and indirect left turns by creating trajectory guidelines from observed data. Other research focuses on behavioral models beyond discrete choice. Gavriilidou et al. (2019b) introduce a game-theoretical model describing cyclists as utility optimizers regarding physical exertion, path deviation, and collision nearness. Hoogendoorn et al. (2021) add risky and cooperative riding to this framework. For automobile drivers, researchers push behavioral models to consider endogenous cognitive processes. van Lint and Calvert (2018) model a driver's situational awareness and capacity to handle an aggregation of mental demand to describe driver distraction. Siebke et al. (2023) model the gaze and cognitive map that drivers build their decisions on and apply this model to virtual conflicts with cyclists. This shows how human perception and behavior modeling can help to simulate unsafe situations.

3.4. Calibration

Calibration finds model parameter values so that the output matches data gathered from the desired scenario. This section reviews how the included studies perform calibration and compares this to established frameworks.

Importantly, model quantities set directly to a desired value and iteratively tuning model parameters so that a resulting model quantity produces the desired values must be distinguished. In the first category, VISSIM allows directly configuring the vehicle flow per link (PTV Group, 2023). Researchers choose values to match observed field data (e.g., Monsere et al., 2019), assumptions about future traffic (e.g., Campisi et al., 2020), or theoretical capacity (e.g., Joo et al., 2012). With the exception of Xu et al. (2023), verification that the simulation achieves the desired output is omitted. Other parameters frequently set in this fashion are speed and acceleration profiles. Studies often describe these steps as calibration, but the missing tuning procedures and goodness-of-fit measurements do not satisfy established definitions as given by Buisson et al. (2014). Rather, these properties can be considered the model input.

Many properties cannot be set directly. In that case, researchers tune parameters so that an output property, or measure of performance, satisfies a goodness-of-fit criterion. To aggregate statistics, researchers repeat simulation runs with different random seeds. Table 2 shows an overview of the parameters and performance measures found in our review. Among these, six tune the behavioral parameters of existing software packages considering safety measures. Bahmankhah et al. (2019b) adjust a conflict threshold to replicate conflict counts observed by a trained professional and report mean absolute percentage errors $< 15\%$. Lemcke et al. (2021) extensively analyze result sensitivity towards behavioral parameters and conflict thresholds in VISSIM and demonstrate how calibration improves conflict number predictions (default: 8; calibrated: 14.4; field-observed: 25). However, their best result still significantly underestimates field observations, and they observe a significant dependency on the random seed. Russo et al. (2022) tune the same parameters, and while three study locations produce good results, six locations are rejected due to unsuccessful calibration. Monsere et al. (2019) manually tune yielding behavior to match field-observed conflict frequencies. Simulated and real conflicts show similar trends, but the overall conflict numbers are small with significant relative errors. Finally, Wu et al. (2014) report a relative error of $< 12\%$ between the simulated and video-observed conflicts without providing details on the calibration procedure. Alecsandru et al. (2010) and Tafidis et al. (2019) report qualitative calibration of the RU behavior by visual inspection.

Regarding new external models, researchers use extensive step-by-step procedures with rich data sources. For example, Langer et al. (2023) present a scheme for calibration based on crash and injury databases. After linking cyclist injuries with crashes through a crash severity model, they calibrate vehicle speeds and a human error model to recreate crash statistics. Several works (e.g., Liu et al., 2020; Ni et al., 2023) extensively calibrate the individual components of their overtaking and turning models by comparing them to real video-captured trajectories and observed maneuver types. While this does not directly tune safety, it controls the interactions responsible for safety. Xu et al. (2023) propose a unique approach where simulated trajectories are compared to trajectories generated by human riders tasked with the simulated scene in a virtual reality cycling simulator.

All reviewed studies attempt to quantify safety, but only the above consider safety during calibration. Others omit calibration or only calibrate travel times, queue lengths, traffic volumes, and trip delays. This does not comply with the standard framework for micro-simulation studies introduced by Buisson et al. (2014). They question the predictive quality of uncalibrated models and deduce that all evaluated performance measures must be calibrated according to measured data. On the other hand, Guo et al. (2019) explicitly analyze calibration for micro-simulated safety and conclude that tuning operational performance measures has a larger influence on the safety results than RU behavior. However, this analysis is limited to vehicular traffic and does not consider the complicated irregular cycling behaviors. For cycling, Lemcke et al. (2021) show that behavioral parameters significantly impact conflicts. Similar sensitivity has been observed by Pauwels et al. (2022) for interactions between cyclists and AVs.

3.5. Validation

Validation examines the predictive quality of the calibrated model. This section reviews validation methods from the included studies and compares them to established frameworks. Twelve studies report validation of their models, with seven validating a safety measure using two approaches.

Three studies feed a different input dataset to their calibrated model and test a safety measure. Liu et al. (2020) validate based on demand data from another road with different characteristics. Their overtaking model achieves an error of $< 31\%$ in predicting over-taking conflicts. Xu et al. (2023) extensively analyze if the conflicts created by their VISSIM-Unity co-simulation correspond to the results produced by human cyclists tasked with the same situations in a virtual reality cycling simulator. They observe that humans create more conflicts, but the differences are not statistically significant ($p = 0.278$). The same conclusion holds for crashes predicted based on the observed conflicts. Finally, Sun et al. (2020) test their VISSIM-trained DNN model for right-turning cars for a small sample of field-observed conflicts with cyclists and obtain a reasonable good conflict severity fit (2.56 s field-observed average PET vs. 2.79 s simulated average PET). Not directly considering safety, Ni et al. (2023) validate their maneuver prediction models and scores up to 91.76% accuracy. The other four studies report validation as the evaluation of performance measures other than the calibration measures, but maintain the same input data.

Sargent describes validation as "determining whether the simulation [...] has the accuracy required for the model's intended purpose over the domain of the model's intended applicability" (Sargent, 2010, p. 174). The purpose of most reviewed studies is to examine the traffic effects of a scenario change. After Buisson et al. (2014), this requires testing the calibrated model's capability to approximate a dataset not used for calibration (predictive validation) with the performance measures of the desired analysis. It ensures the predictive quality in response to input changes. Regarding safety, the validations of Liu et al. (2020), Xu et al. (2023), and Sun et al. (2020) conform to this

Table 2
Parameters and performance measures of the tuning-based calibration procedures in the reviewed studies. This does not show quantities set directly to desired values (i.e., input parameters).

Reference	Parameters	Performance measures			Tuning	Goodness of Fit
		Operational	Behavioral	Safety		
Bahmankhah and Coelho (2017)	–	saturation flow*			–	GEH, MAPE, R^2
Bahmankhah et al. (2019a)	car-following*, lane-change*, lateral behavior*, max. deceleration*, visibility range, safety gaps, TTC threshold	saturation flow*			–	GEH, MAPE, R^2
Bahmankhah et al. (2019b)	W74 car-following*, lane-change*, lateral behavior*, max. deceleration*, visibility range*, safety distances*	saturation flow[c,b], travel time[c,b]		conflicts[c-b]	sensitivity analysis	GEH, MAPE, R^2
Guhathakurta et al. (2023)	–	travel time	speed,		–	absolute comparison
Li et al. (2020)	lane-change rules[c]	maneuver type[b]			linear regression, logistic regression	Likelihood-ratio test, Wald-test, R^2
Liu et al. (2020)	sublane width, sublane choice[c,b], overtaking motivation[b], gap acceptance[b], car-following[b]	travel time[b]	lateral position[b], gap acceptance[b]		genetic algorithm, maximum likelihood	MAE, likelihood
Langer et al. (2023)	perception errors[c,b], reaction time[c,b], desired speed[c]		speed[c]	crash rates[c-c, c-b]	genetic algorithm, iterative proportional fitting	KS
Lemcke et al. (2021)	W74 car-following[c,b], lack of attention[c,b], safety distances[c,b]	delay*, platoon ratio*		conflicts*	manual	numerical absolute comparison
Monsere et al. (2019)	desired speed[c,b], priority rules[c,b]			conflicts[c-c, c-b]	manual	graphical absolute comparison
Ni et al. (2023)	Comfort zone model[b], Bayesian decision network[b], Social force model[b], car-following[b]		comfort zone size[b], maneuver type[b], speed[b]		regression analysis, K2-Algorithm, EM-algorithm, genetic algorithm	SSE, F-Test, RE, RMSPE
Preston and Pulugurtha (2021)	W74 car-following[c], lane-change[c], speed distributions[c]	number of vehicles[c], queue length*	approach speeds[b,c]		manual	RE, visual inspection
Ren et al. (2016)	cellular automata[b]	clearance time[b]			–	RE
Russo et al. (2022)	W74 car-following*, safety distances*	queue lengths*		conflicts[b-c]	–	numerical absolute comparison
Silva et al. (2023)	W74 car-following[c]	crossing time[b,c]			–	t-test
Sun et al. (2020)	DNN trajectory generator[c]		trajectory[c]		backpropagation	MSE
Wu et al. (2014)	–	delay[b,c]		conflicts[b,c]	–	RE
Xu et al. (2023)	W99 car-following[b], Unity path-finding[b]		lateral position[b], speed[b]		regression analysis, manual tuning	R^2, visual comparison of trajectories

(*) traffic mode of the tuned parameters or analyzed performance measure not reported.
([c]) heavy vehicle or (automated) car model parameters and performance measures.
([b]) bicycle model parameters and performance measures.
(–) no information.
Initialisms: Deep Neural Network (DNN), Geoffrey E. Havers Statistic (GEH), Kolmogorov–Smirnov test (KS), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), Relative Error (RE), Root Mean Squared Percentage Error (RMSPE), regression coefficient of determination (R^2), Sum of Squared Errors (SSE).

definition. Without input variation, the other reviewed studies do not explore how the model predicts traffic performance under different conditions. In some cases, prediction may not be the purpose of a study. For example, Lemcke et al. (2021) investigate the relationship between simulated and field-observed conflicts on one intersection. The model input does not vary between calibration and experimentation. After Sargent (2010), close examination and comparison of the model behavior constitutes operational validation. It analyses the calibration outcome by means ranging from different performance measures to sensitivity analysis. To prevent confusion with predictive validation, we call this process verification from now on. Verification does not build confidence into a typical micro-simulation's primary purpose to generalize over varying conditions. For example, Preston and Pulugurtha (2021) experiment with different intersection layouts. Their verification considers queue lengths on an existing layout but does not show the model's predictive quality regarding the alternative layout. This issue is inherent in studies simulating future scenarios where data collection is impossible.

3.6. Safety assessment

This section reviews how the included studies quantify cycling safety and compares this to the general traffic safety domain. We identified three categories of safety assessment. Most studies (28 of 42) are based on crash surrogacy. Eight studies report simulated collisions. Six studies measure safety through exposure to risk factors.

Risk exposure

Hübner et al. (2017) use field data to identify high-risk areas in Berlin. They measure the safety of simulated routes by time spent in these areas and their count. Alecsandru et al. (2010) and Wu et al. (2014) define conflict zones in intersecting traffic streams. The number of RUs in the zone describes exposure used to evaluate signal control strategies. Similarly, Thompson et al. (2015, 2017) count potential collisions based on RUs simultaneously present on the same network node. Wallentin and Loidl (2016) successfully fit a linear regression model linking simulated encounters with accident reports.

In these studies, safety is not estimated based on interactions and event causality but by exposure alone. Vanparijs et al. (2015) highlight the significance of exposure on quantitative cycling safety analysis. Only controlling for exposure allows meaningful comparison. However, the outcome of an event is also influenced by additional factors. Hence, measuring exposure alone can give safety indications but not the whole picture.

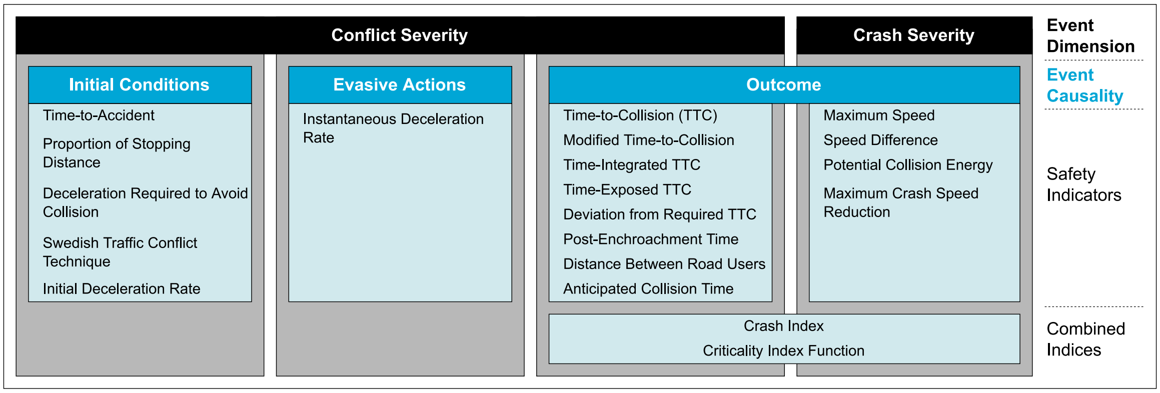


Fig. 3. Typology of surrogate safety indicators used in the reviewed studies.

Conflicts as collision surrogates

Surrogate safety assessment analyses individual RU interactions. Its theoretical foundation is the relationship between high-frequency zero-risk interactions and low-frequency crashes postulated by Hydén (1987). In between are traffic conflicts with an increasing severity or closeness to a collision. Observing frequently occurring conflicts enables estimating rare collisions. The nature of the crash-conflict relationship is debated, making it difficult to predict incident risk numbers conclusively (Arun et al., 2021). Surrogate Measures of Safety (SMoS) characterize RU interactions to identify conflicts.

Two SMoS dominate our review. Eighteen studies calculate the Post-Encroachment Time (PET) to identify conflicts and describe severity. It measures the time between a RU leaving a conflict area and a second one entering it (Allen et al., 1978). Similarly, twenty studies use the TTC. It describes the remaining time to a collision for RUs on a collision path assuming constant speed and heading (Hayward, 1972).

Fig. 3 shows a typology of safety indicators in our review that combines event severity and event dimension. The event dimension introduced by Arun et al. (2021) distinguishes indicators that quantify nearness to a collision (conflict severity after Hydén, 1987) and indicators that quantify the injury risk of a hypothetical collision resulting from the conflict (crash severity). Johnsson et al. (2018) propose to sort SMoS by event causality. In the causal conflict model by Davis et al. (2011), initial conditions and evasive actions define the outcome of an event. TTC and PET are indicators of the severity of a conflict outcome. An example of an indicator for the initial conflict condition severity is the Time-To-Accident, the TTC at the start of an evasive maneuver (Hydén, 1987), used by Pauwels et al. (2022). The instantaneous deceleration used by Li et al. (2011) is the only indicator of evasive action severity in our review. For crash severity, researchers report, among others, the Maximum Speed or Speed Difference. Finally, combined indicators measure the severity of the conflict outcome and the crash severity. Pauwels et al. (2022) use the Crash Index, which fuses a TTC-based measure with speed. To detect conflicts with these indicators, researchers apply thresholds. If an interaction exceeds the threshold, a conflict is registered. The magnitude of the SMoS describes its severity. Gettman et al. (2008) automate the detection and evaluation with their SSAM, which 16 of 20 VISSIM studies employ. In our review, only Xu et al. (2023) use an extreme value theory approach to predict concrete crash risk based on the observed conflicts.

Several reviewed studies report issues with surrogate safety assessment. Some authors encounter virtual crashes when the model's behavior parameters are tuned to resemble realistic cycling. The treatments of these collisions vary. Monsere et al. (2019) include collisions in their assessment but do not report the severity based on TTC or PET as they consider them unreliable. Tafidis et al. (2019) ignore collisions by applying a lower TTC threshold. Bahmankhah et al. (2019b) recalibrate until collisions account for less than 10 % of all conflicts. The creators of SSAM acknowledge this problem (Gettman et al., 2008). As most microscopic simulations are intended to be collision-free, they conclude that these events do not constitute inaccurate measurements but shortcomings in the simulation models. Until models can be improved, they recommend recalibration or spatial filters that limit safety assessment to areas with accurate behavior. Huang et al. (2013) show that many collisions at intersections can be resolved through small changes to the road topology. While this issue was already present for cars, the difficulty of tuning lane-based models to complicated cyclist maneuvers could exacerbate its severity. Additionally, Russo et al. (2022) observe that SSAM does not detect conflicts for some perpendicular interactions, although they were realistically simulated based on visual assessment. They did not investigate this error further. A potential explanation is that SSAM only detects conflicts that simultaneously satisfy TTC and PET thresholds. Correspondingly, the SSAM validation study (Gettman et al., 2008) observes scenarios where conflicts are missed because a vehicle brakes abruptly for a crossing RU. These situations lead to small TTCs, but waiting before continuing to drive can create large PETs and conflicts are not registered. The resulting bias could be especially severe for cycling due to complex and diverse interactions.

In their validation of TTC and PET based on real-world conflicts between cyclists and motor vehicles, Johnsson et al. (2021) establish a significant correlation between conflict frequency and recorded crashes. However, this was weaker than the correlation between crashes and exposure. Corrected for exposure, PET does not correctly rank the safety over multiple locations. Furthermore, TTC results in many false positives, where visual examination did not confirm the conflict severity indicated by the temporal proximity. They conclude that speed-based indicators should supplement indicators based on temporal proximity.

Simulated collisions

Eight studies simulate collisions and count their frequencies. In Thompson et al. (2016, 2020), a collision occurs based on a probabilistic model of RUs association with each other. Denk et al. (2022) sample collisions from parallel stochastic models of avoidance measures. Ren et al. (2023) and Pechinger et al. (2021, 2023) count the collisions that their AV control algorithms create. Karkhanis et al. (2020) deactivate yielding and count events where RUs go below a distance threshold as collisions. Langer et al. (2023) introduce perception errors to recreate real-world injury and incident statistics. A high-level database provides police-reported injuries from four years at the study location. A low-level database lists incidents from two other cities with detailed crash types. They map crash types to similar locations on intersections in the study location and calibrate the simulation to reproduce the crash statistics on both levels. While they report promising results for high traffic volumes, small volumes do not produce significant numbers. This highlights that crashes are inherently rare, and observing them in simulation and naturalistic data requires long observation periods. In Langer et al. (2023), the model reproduces the real-world statistics of four years in 60 simulated days, meaning simulated interactions are riskier than in real life. To overcome the limitation of long runtimes for long simulation periods but still enable large sample numbers, Denk et al. (2022) propose a sampling-based approach. For right-hook conflicts, they first generate undisturbed trajectories. Random sampling and overlapping undisturbed trajectories creates initial conditions for encounters, where the number of generated encounters is based on the expected exposure. Sampling from parallel stochastic models of driver perception, cyclist behavior, and a V2X automated emergency braking system determines if an encounter results in a collision. While this approach captures some aspects of crash causality, the simplifications made in the stochastic models do not fully capture the interaction process and consequently underestimate real collision frequencies.

Event causality in simulated conflicts and collisions

For meaningful simulated collisions, the causal chain of events needs to be realistically represented. Following Davis et al. (2011), this includes the initial conditions, evasive actions, and the event outcome. The capability of current models to describe exposure and some aspects of human error may create initial conditions that precede unsafe events, but lane-based models constrain possible evasive actions. Hence, the outcome may not sufficiently correlate with reality. The prevalence of event outcome indicators may lead to unrealistic results, even if studies measure conflicts as collision surrogates. Johnsson et al. (2018) propose to use multiple indicators that represent all categories of the typology (Fig. 3) and to limit measurements to initial conditions if the other steps are insufficiently modeled. However, in light of complex and diverse cycling behaviors, the existing simulation packages often already have cycling-specific shortcomings regarding the initial conditions.

3.7. Scenarios and interactions

The included studies differ in the spatial extent of their scenario. While some studies simulate metropolitan areas (e.g., Hübner et al., 2017), others focus on limited neighborhoods (e.g., Tafidis et al., 2019), intersections (e.g., Russo et al., 2022) or road segments (e.g., Liu et al., 2020).

Safety is a result of individual interactions, so scenario-specific interaction types significantly affect the results. However, among the studies using existing models, most do not report their results on the interaction level and only analyze aggregated measurements. Below, we summarize available comments on model performance for specific interactions with cars and between cyclists.

Crossing interactions

A frequent type is the encroachment of turning vehicles on straight-going cyclists. Fundamentally, the intersection models in existing software packages can simulate priority rules for encroaching traffic. Lemcke et al. (2021) find VISSIM settings that qualitatively enable the correct behavior of motorists crossing a cycling path to access their turn-lane. Similarly, Russo et al. (2022) qualitatively observe the expected conflict types in right-turn scenarios. Quantitatively, conflict counts do not match field observations in both cases. Monsere et al. (2019) confirm this by counting unrealistically many car-bicycle conflicts in a protected intersection design. While they suspect issues in managing isolated traffic streams, they also choose aggressive behavioral VISSIM settings to prevent defensive yielding. For SUMO, the documentation warns about unrealistic yielding of cyclists because they use the same safety gaps as cars (German Aerospace Center, 2023a). Grigoropoulos et al. (2022) successfully calibrate VISSIM to field-observed cyclist queue discharge times and automobile waiting times in right-hook conflicts, showing that VISSIM successfully models important operational characteristics on a group level. Considering the issues with interaction-level behavior reported above, this does not guarantee acceptable safety performance.

Outside the included studies, Ma et al. (2017) show that VISSIM-generated trajectories of left-turning automobiles only capture a small fraction of the spatial diversity of field observations and thus cannot represent realistic interaction behaviors. They develop a spatially continuous 2D model that better predicts the variety of car maneuvers in left-turning conflicts with cyclists. After calibration, their model achieves a good fit between average simulated (3.94 s) and field-observed (3.99 s) PET. Other new models for the same scenario show similar performance (Ni et al., 2019; Sun et al., 2020).

Interactions on shared lanes

In scenarios with shared infrastructure, RUs interact longitudinally while following and laterally while overtaking, changing lanes, or mixing in open spaces. The default lane-based models (see Section 3.3) offer parameters to adapt longitudinal dynamics to real-world observations. Authors generally do not report issues regarding longitudinal dynamics. For lateral interactions, Monsere et al. (2019) analyze bicycle-vehicle conflicts in a shared right-turn lane and observe that the VISSIM lane-based model does not sufficiently use the lateral space. In contrast, Russo et al. (2022) achieve credible conflict numbers for their shared right-turn lane. Additionally, Sun et al. (2019) analyze the VISSIM overtaking model on two-lane two-way roads with cyclists and observe unrealistically weak reactions of overtaking RUs to oncoming traffic.

Authors develop new models outside existing simulation packages to improve interactions in shared lanes and spaces. In turning conflicts, more complex rule-based decision-making increases the plausibility of interactions with oncoming traffic (Sun et al., 2019). Still, calibration attempts reveal too little diversity of the simulated travel times and spatial trajectory distributions compared to field observations. Other studies create models for overtaking cyclists on one-way roads that intrude into vehicle lanes. The sub-lane model of Liu et al. (2020) achieves realistic travel times but underestimates overtaking events due to the limited sub-lane resolution. Additionally, their lateral position model does not generalize well for different calibration and validation site characteristics. For the same scenario, the residual overtaking maneuver prediction error of the cellular automata approach of Li et al. (2020) after calibration ranges between 3.5% and 13%. The social-force approach of Ni et al. (2023) correctly predicts > 91% of overtaking events in the calibration dataset and shows promising results for lateral distance, trajectory distribution, travel time and safety. However, the latter two studies did not perform a full predictive validation. Ren et al. (2016) specifically address the lateral dispersion effect of straight-going cyclists on intersections with a cellular automata approach. For roundabouts, trajectories created by the VISSIM-Unity co-simulation of Xu et al. (2023) graphically fit human trajectories well. Still, the simulation produces a smaller headway than human riders, creating fewer conflicts and crashes. Additionally, they observe differences in lateral placement on the lane without influence on conflict occurrence.

Specialized cycling infrastructure

Researchers simulate specialized bicycle infrastructure like bike boxes. As simulation packages often lack native support, researchers attempt workarounds. In VISSIM, Russo et al. (2022) extend a single lane into the intersection as a “bike box”. They do not measure realistic conflicts. In contrast, Joo et al. (2012) implement a bike box through adjacent narrow lanes with bicycle priority and do not report issues. For SUMO, Grigoropoulos et al. (2019) guide how to model advisory bicycle lanes, bicycle boxes, and indirect left turns within the package limitations. In a later study, Grigoropoulos et al. (2022) conclude that simulated cyclist behavior around bike-boxes qualitatively aligns well with naturalistic traffic observations, but results were not numerically validated.

Variability in road user behavior

Monsere et al. (2019) and Russo et al. (2022) attempt to calibrate a single parameter set for multiple intersections and observe that not all local phenomena could be captured. This may be caused by the single parameter set, the lane-based model, or the behavioral rule set. Even for single scenarios, Lemcke et al. (2021) report that their calibrated model still underestimates field-observed conflicts, pointing to model shortcomings for safety-relevant behaviors. More positive reports of bicycle model performance, like the COWI manual (COWI, 2013) or the W99 calibration efforts for cyclists by Kaths et al. (2021), target operational characteristics without analyzing safety-critical behavior. But even here, Kaths et al. (2021) identify lacking options to assign distributions to car-following parameters as a limiting factor and highlight the significance of lateral motion to realistic cyclist behavior. For simulated and observed pedestrian-automobile crosswalk conflicts, Wu et al. (2018) show that illegal behaviors contribute significantly to the residual. Guhathakurta et al. (2023) expect an underestimation of cycling conflicts due to the same effect but argue that a delta comparison limited to the modeled factors is possible. Qualitatively, more researchers from our review find the models in existing packages insufficient to represent complex cycling behavior (Ledezma-Navarro et al., 2018; Xu et al., 2023). Thompson et al. (2020) point out missing knowledge of incident causality as a critical shortcoming.

These insights make it impossible to single out (un-)suitable scenarios. Current simulation packages fundamentally enable the creation of many infrastructure layouts and interaction types. However, the summarized issues highlight a lack of robustness to unsuitable configurations and a lack of validation for specific scenarios. Unfortunately, many studies in our review analyze safety based on microscopic interactions simulated by the existing software packages but do not validate, verify, or comment on whether the simulation functions correctly. As a result, a simulation might seem plausible macroscopically but produce unrealistic microscopic interactions. Improved external cyclist models can provide solutions for the shortcomings of the default models but are scenario-specific and mostly yet to be comprehensively validated.

4. Framework for simulated cycling safety assessment

After reviewing existing studies, we present a methodological framework for simulated cycling safety assessment (Fig. 4). The framework inherits building blocks from the Barceló (2010) conceptual framework of micro-simulation models and the calibration guidelines of Buisson et al. (2014). Many fundamental aspects outlined in the guidelines for microscopic simulation by the US Federal Highway Administration Traffic Analysis Toolbox Volume III (TAT) for assessing operational characteristics (Wunderlich et al., 2019) also apply to cycling safety. We reiterate those only where related issues became apparent in the reviewed application studies. Instead, we focus on the additional cycling and safety assessment requirements and highlight differences. We take up the suggestion of virtual Randomized Controlled Trials (vRCTs) proposed by Brunner et al. (2019) and used by Denk et al. (2022). They apply the structured and established procedure of randomized controlled trials in medicine to simulated interventions in traffic systems, arguing that the considerable overall complexity of a traffic system requires similar statistical tests as investigations of the human body. This may help to systematically address the stochasticity of traffic simulation and create robust results.

Our review raises further concerns regarding the capability of current micro-simulation packages to simulate cycling safety. Calibration attempts have high residuals, and no application study comprehensively validates their model for cycling safety. Those studies looking at individual interactions report limitations of the default lane-based models to capture the diversity of movement patterns. While perception errors and rule noncompliance can be simulated to create conflicts, the RU behavior in critical conflict situations is usually not addressed. Consequently, conflict causality is not sufficiently reflected in the existing models. Automated conflict-based assessment may miss unsafe situations due to unsuitable indicator combinations. In summary, we do not find sufficient evidence that the currently available simulation packages enable cycling safety assessment with the same predictive quality as the operational characteristics of car traffic. Hence, researchers may not rely on current tools as validated out-of-the-box solutions. Instead, one must carefully analyze the requirements of study conditions, individually develop, calibrate, and validate model components and communicate underlying assumptions and limitations. Examples of this process can be found in the reviewed studies that create new individual models. Our framework aids this process by listing necessary steps and highlighting methodological requirements.

We discuss the building blocks of the framework and their methodological requirements in the following subsection. Then, we discuss the framework's application potential and summarize the research gaps.

4.1. Methodological requirements

In our framework, every step defines requirements for the building blocks of the following steps.

Problem definition

During problem definition, researchers must describe the study scenario. Following Buisson et al. (2014), the factors influencing the traffic scenario and cycling safety must be identified based on field observations or related literature. Researchers may limit the selection of factors to those relevant to the research questions. The suitability of micro-simulation to answer a research question depends on the validated capability of the model to describe the relevant factors. With safety being a characteristic of RU interactions, researchers must pay special attention to behavioral patterns. Simulations of a real-world scenario are most suitable as the characteristics of the specific locations may be observed and used to design, calibrate, and validate the model. Generic and predictive scenarios without an immediate real-world counterpart are especially valuable for research questions on future traffic developments but require research into new models with proven intrinsic validity. For statistical testing with vRCT after Denk et al. (2022), the formulation of research hypotheses, as well as a baseline (i.e., "control group") without the proposed intervention, is required.

Environment selection

Based on the critical factors of cycling safety defined in the previous step, researchers have to select the building blocks of the simulation. Specifically, the simulation environment and its bicycle models need to be capable of modeling the RU behavior, relevant infrastructure elements and their geometry, as well as traffic characteristics like travel speeds, queuing patterns, and stop waves.

Our reviewed studies successfully model bicycle flow and traffic control using established tools (e.g., Russo et al., 2022; Xu et al., 2023; Bahmankhah and Coelho, 2017). However, our review reveals significant shortcomings of the default lane-based models for cycling safety. In scenarios with one dominant direction of travel and without excessive lateral dynamics, studies may continue the work of Lemcke et al. (2021) to explore the validity of the lane-based approach further. For more complex irregular interactions, this review points towards the unsuitability of the default lane-based approach. Extending the TAT, this puts a focus on selecting and modifying or newly developing a traffic model that can describe the specific required behaviors and interactions. Our review includes several examples of new models for specific scenarios that significantly improve simulated road user behavior (e.g., Ni et al., 2023; Xu et al., 2023). Similarly, infrastructure models must be able to accommodate cycling behaviors. While the main elements are available in most environments, restrictions apply for special cycling infrastructure (e.g., bicycle boxes).

As exposure is a prerequisite for (un)safe interactions, performance measures should both test operational and safety performance. This

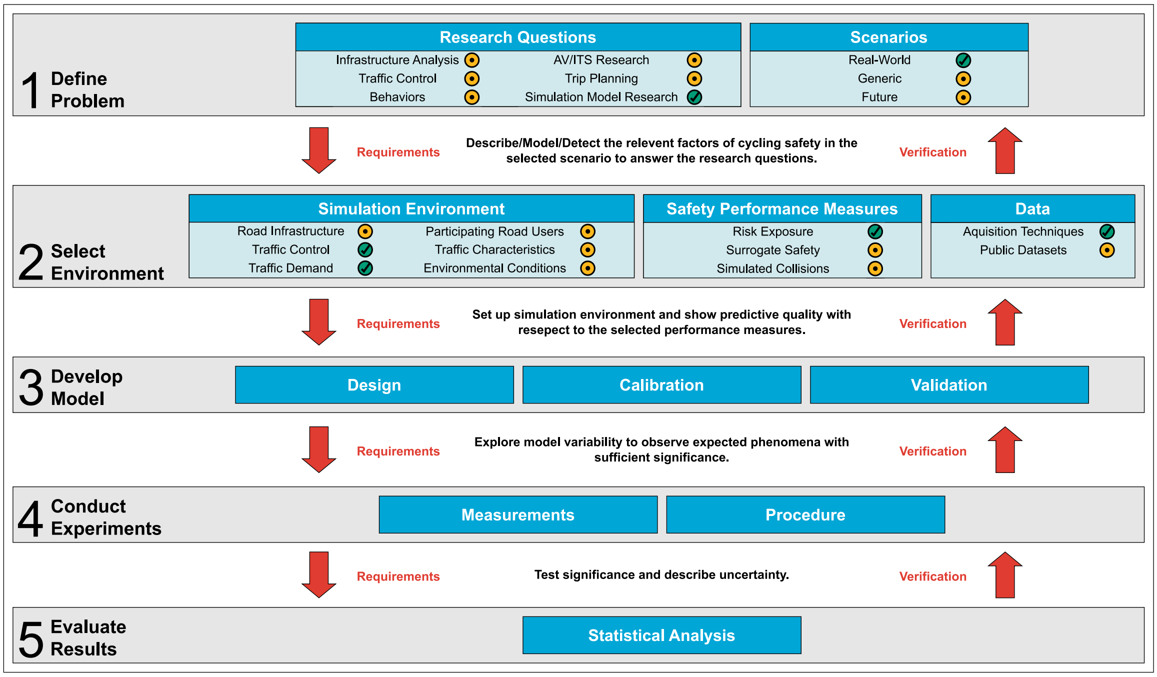


Fig. 4. Methodical framework for cycling safety assessment with micro-simulation. Green ticks indicate model components ready for use in cycling studies. Yellow dots indicate components that require further scenario-specific research.

extends the TAT requirements for a local and a global operational measure by at least one safety performance measure that must be chosen to answer the research question and reliably detect the expected unsafe situations. For interaction level assessments, SMoS measure different varieties of hazardous situations and are not necessarily interchangeable. Johnsson et al. (2018) provide three recommendations for choosing SMoS for cycling: Firstly, a combination of multiple indicators is necessary to detect unsafe cycling. This should include indicators for initial conditions, evasive actions, and conflict outcomes, provided the RU models describe all parts. Secondly, indicators solely based on temporal proximity may overestimate conflict frequency, and a combination with speed-based measures should be evaluated. Thirdly, crash severity is essential for cyclists because of their high vulnerability compared to car passengers.

SSAM should be used with caution. Firstly, it only measures the conflict outcome, and secondly, the conflict detection based on combined TTC and PET thresholds can miss dangerous situations. Cycling creates complex dynamic behaviors that may be especially hard to capture correctly. This much increases the requirement for visual inspection of the simulation outcome compared to the TAT, not only for verification of the simulation but also to understand what situations the safety indicators detect. During our review, building confidence in the simulation quality was frequently limited by the lack of dissemination of spatiotemporal results. To overcome this, researchers should more often show example trajectories, trajectory distributions, or animations of their simulations.

If safety assessment is based on risk exposure measures alone, this requires prior knowledge of the associated risk and neglects crash causality. Thus, it may provide an incomplete picture. Still, as exposure is a prerequisite to crashes, Wallentin and Loidl (2016) show that it reveals a fundamental insight into safety. With existing models providing realistic traffic flows, exposure measurements are possible if developing an improved model is out of scope. On the other end of the spectrum, simulated crashes must consider event causality and, hence, are only a viable measure if this can be simulated sufficiently. In all cases, predicting a concrete crash risk is currently limited to particular scenarios with realistic models throughout the simulated causality chain.

Lastly, calibration and validation data must be collected. It must showcase the relevant factors influencing cycling safety and enable the extraction of the performance measures.

Model development

Based on the selected environment, the traffic model must be developed. In addition to the conventional setup of road geometry, traffic control, and demand, necessary external behavioral models must be integrated. New models, complex behavior depending on local context, and the limited capabilities of established models increase the focus on model calibration and, compared to the TAT, require validation.

Calibration and validation must target the specific requirements of the study objective, as defined by the environment and research questions. After Buisson et al. (2014), it must test all performance measures intended for experimentation. If the experiment includes model input changes (e.g., different road layouts, traffic volumes, or mode share), the model's capability to predict the safety outcome of these changes must be validated. If previous works have calibrated and validated sub-models in comparable conditions, researchers may rely on this validity (Buisson et al., 2014). However, our review reveals the scarcity of successful cyclist model validations for micro-simulation and highlights their instability under different scenarios. Hence, researchers must perform their own calibration and validation in most cases. Without it, the simulation results might not be informative about real-world performance. In future scenarios like mixed traffic of cyclists and AVs, data-driven validation may be impossible. Here, further dissemination of the underlying processes may create sub-models that can be validated with data, are based on physical principles, or enable strong assumptions. When validation is impossible, studies must communicate their hypothetical character and underlying assumptions.

To ease the calibration complexity, researchers typically apply multi-stage pipelines that first tune operational characteristics and then focus on safety (Huang et al., 2013; Guo et al., 2019). For the calibration of operational performance, the TAT does not require multiple runs, assuming little dependency on stochastic behaviors. In contrast, microscopic RU interactions depend significantly on the distribution of behavioral parameters. Hence, safety calibration and validation require multiple runs with different random seeds.

Experimentation

During experimentation, researchers run the simulation and record performance measures. The traffic model defines the requirements for the experimental procedure. A simulation run is characterized by its run time, the sampling rate, and a warm-up period to populate the network. The sampling rate must allow sufficient temporal resolution to generate realistic interactions. As traffic simulations include random processes, a simulation has to be run multiple times. Reviewed studies report results significantly depending on the random seed of a simulation run (Lemcke et al., 2021; Russo et al., 2022). Large fluctuations between runs indicate that a model contains much randomness and that experiments require more runs to explore the parameter space. Hollander and Liu (2008) describe how a sufficient number of runs may be derived to estimate the mean of a performance measure reliably. The TAT recommends four runs for this initial estimation. For safety, Denk et al. (2022) highlight that critical situations may arise from rare behaviors at the tails of parameter distributions. More than four runs may be necessary to observe these for the first mean and standard deviation estimation. The vRCT framework calculates the required number of encounters based on the desired statistical significance level, effect size, and statistical power.

Evaluation

Measurements of a random variable in the experimentation step require researchers to report measurement statistics. If scenarios are compared, the comparison must be analyzed for statistical significance. Here, our review did not identify unique aspects of cycling safety.

4.2. Applications and outlook

We intend our framework to be a study guideline for researchers and practitioners. However, the shortcomings of current cycling simulations are significant, and the framework formulates very high requirements. These can be met with reasonable effort for limited research questions relying on risk exposure or scenarios with a dominant direction of motion. In many cases, however, the need for behavioral research, development of new traffic models, data collection, and calibration/validation might exceed the scope of an application study. Analyzing the requirements for a specific study with our framework will, however, reveal research opportunities that might lead to the development of more comprehensive simulations.

The various applications we observed, including emerging challenges of our future transport systems, represent a clear need for integrated simulation environments. Creating cycling-friendly transport systems and introducing AVs, potentially at the same time, raises a multitude of safety-related research questions regarding infrastructure, operations, and logistics applications. Elevating microscopic simulation to a tool that comprehensively answers these questions could help development and boost safe real-world implementations. Further disseminating the underlying behavioral and mechanical processes to create sub-microscopic models is a promising approach. For example, models of driver distraction (van Lint and Calvert, 2018) may similarly apply to cyclist-driver interactions. The agent-based approach of modeling every RU with autonomous intelligence can put a focus on individual behaviors and motivations (Nguyen et al., 2021). For vehicle dynamics, Pechinger et al. (2021, 2023) couple micro-simulations with existing vehicle dynamics models and Schmidt et al. (2024) show that this may also be applied to bicycles. More research needs to focus on cycling to create integrated simulation environments for reliable predictions on multiple levels of abstraction.

4.3. Research agenda

Our framework highlights that significant research is required to unlock micro-simulation as a cycling safety tool and create integrated traffic simulations of future transportation.

Cyclist Behavior: To develop and calibrate micro-simulation models, cyclists' typical behavior in conflicts must be known. While research into this is active, many specific scenarios are insufficiently investigated. Specifically, developing predictive safety models needs information on RU behavior in severe conflict situations at the final stages of crash causality. Research on a global level is necessary to understand diverse traffic conditions around the world.

Conflict Causality Models: To make simulated conflicts meaningful, sub-models for the underlying processes of cycling behavior are needed. Research may build on existing frameworks like the social force concept, task-demand models, or human learning to focus on weakly disciplined non-lane-based behavior and human factors. Integrating these sub-models into a single framework would enable safety analysis on the interaction and system levels.

Validation of Existing Models: Comprehensive model validations considering cycling safety are scarce. Determining under which conditions existing and new approaches produce realistic results would help researchers and practitioners identify appropriate models for a specific scenario. Here, research into model transferability and generalizability is needed.

Cycling Field Data: With data availability limiting model development and validations across scenarios, more research should make their data available to others. To this end, creating a comprehensive repository of cycling data would greatly benefit the collective effort to create better models and more reliable application studies.

5. Conclusions

Due to its practical benefits, researchers apply microscopic traffic simulation to assess cycling safety in current and future traffic scenarios. However, previous research has voiced concerns about whether the underlying models can predict safety-relevant behavior. We analyzed the methodology of 42 studies quantifying micro-simulated cycling safety to answer the following questions.

***RQ**1: Can micro-simulation model the safety-relevant behavior and riding dynamics of cyclists?* Studies predominantly tune the lane-based models of existing simulation packages for automobile traffic to resemble cycling. Parameters affecting gap acceptance, visibility, and attention are set to create unsafe situations between cars and bicycles. However, detailed calibration attempts show that the simulation underestimates field-observed conflicts (Lemcke et al., 2021). Additionally, researchers observe shortcomings regarding dedicated cycling infrastructure and report difficulties finding broadly applicable model parameters (Monsere et al., 2019; Russo et al., 2022). Cyclists and conflicting cars overly stick to lanes instead of showing variable trajectories (Ma et al., 2017; Roosta et al., 2023), indicating that simulations may miss safety-relevant situations. Our review does not find definitive, successful validations of existing default models regarding the safety of simulated behaviors and interactions. Hence, studies must carefully develop and validate scenario-specific models to achieve reliable results. Research on overtaking models (Ni et al., 2023), perception models (Langer et al., 2023), and turning models (Sun et al., 2020) achieves promising results. It demonstrates how modeling the underlying interaction processes among cyclists and with cars may enable the simulation of the whole chain of conflict causality. In the general domain, spatially fully continuous models (Kaths, 2023) and cognitive models (van Lint and Calvert, 2018) are promising directions to increase simulated event causality further.

***RQ2**: Can micro-simulation models be calibrated and validated to predict traffic conflicts involving cyclists?* Researchers generally calibrate and validate the existing simulation packages for operational characteristics like flow, queue lengths, or trip duration. No application study in our review calibrates and validates safety performance. Further, they do not comment on whether the simulation environment creates realistic behaviors and interactions. This conflicts with established simulation literature demanding that the calibrated and validated performance measures must be identical with experimentation and that distinct datasets must be used for a model to prove its predictive power (Buisson et al., 2014). With a lack of successful validation studies concerning simulated cycling safety, studies may not rely on the general validity of currently available tools. New external models in our review are usually extensively calibrated and show promising performance, but predictive safety validations are scarce. Xu et al. (2023) propose a promising approach for calibration and validation based on human-in-the-loop simulations that needs further analysis for the validity of human behavior in virtual reality.

***RQ3**: Can cycling safety be evaluated based on micro-simulation?* Most studies employ crash surrogacy for safety assessment, mainly using SSAM. This leads to concerns about the completeness of safety results, as SSAM may miss unsafe situations due to serial PET and TTC thresholds. The proposal of Johnsson et al. (2018) to choose safety indicators that cover all relevant parts of conflict causality in the scenario may help to overcome this. However, current traffic models generally do not simulate realistic cycling behaviors through the full conflict process. As a consequence, researchers typically filter simulated collisions. Future models must consider the processes leading to conflicts and behavior in conflicts to make simulated collisions meaningful. In the meantime, Wallentin and Loidl (2016) show that exposure measurements may give limited insights into simulated safety. Reliable conclusions must consider the model stochasticity and employ statistical tools to derive a procedure for experimentation, evaluation and reporting of the results. Denk et al. (2022) demonstrate how the medical gold standard for randomized controlled trials can be applied to simulated traffic assessment.

In summary, current default micro-simulations are not ready for cycling safety assessment. Instead, many building blocks are active research topics. Improved models already exist, but more comprehensive validations are necessary to promote their transfer into micro-simulation applications. Additionally, the application studies of our review often do not follow established best practices. To achieve reliability, researchers and practitioners must carefully determine the factors and behaviors influencing cycling safety specific to the scenario. Then, they must select or create models that can simulate these phenomena and perform calibration and validation regarding safety measures. We present a methodological framework to guide this process.

The framework places high requirements on application and case studies. Traffic simulation is no shortcut to quick results on cycling safety. Overcoming its inherent limitations requires rigor and detailed analysis. The reward is the possibility of making predictions without the technological and ethical limitations of real-world studies. Several reviewed works focus on safety-critical AV functions or connected systems. However, predictions on transportation futures based on models with unclear reliability become weak at best and dangerous at worst. Insufficiently founded safety claims for vulnerable RUs might incentivize developments and policies that eventually create hostile environments and endanger real people.

This review could draw only limited insights from some of the included studies because of insufficient documentation of the method, not beyond high-level facts about the adopted simulation package, the study location, or selected settings. This further highlights the need for a rigorous methodological framework.

Future work must develop comprehensive micro-simulations to aid the design and implementation of safe technological, operational, and logistical innovations for future transportation systems. Especially for AVs, mixed traffic with cyclists is a bottleneck for large-scale adoption. To alleviate the burdens on case studies, research must focus on an improved representation of cycling behaviors and RU interactions. Existing approaches like the social force concept or models of perception, attention, and workload may help to model the processes leading to conflicts and thus enable in-depth analyses and justifiable predictions leading to safety. Incorporated into existing microscopic frameworks, this could answer some of the most pressing future mobility and transportation questions.

CRediT authorship contribution statement

Christoph M. Konrad: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Azita Dabiri:** Writing – review & editing, Validation, Supervision, Conceptualization. **Frederik Schulte:** Writing – review & editing, Validation, Supervision, Conceptualization. **Jason K. Moore:** Writing – review & editing, Validation, Supervision, Conceptualization. **Riender Happee:** Writing – review & editing, Validation, Supervision, Conceptualization.

Declaration of Generative AI in Scientific Writing

During the preparation of this work, the authors used Grammarly to improve spelling, grammar, and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

This research was funded internally by the Faculty of Mechanical Engineering, Delft University of Technology.

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.

Data availability

No data was used for the research described in the article.

References

Aarts, L., Broek, L.v.d., Oude Mulders, J., Decae, R., Bos, N., Goldenbeld, C., Wijlhuizen, G., Mons, C., Hoekstra, A., Bijleveld, F., 2022. De Staat Van De Verkeersveiligheid 2022: Trend in Aantal Verkeersdoden En -Gewonden Daalt Niet [The State of Road Safety 2022: the Trend in the Numbers of Traffic Deaths and Injuries Does Not Decrease]. Rep. R-2022-10, SWOV Institute for Road Safety Research, p. 24, URL: https://swov.nl/en/publicatie/de-staat-van-de-verkeersveiligheid-2022.

Alecsandru, C., Rouhieh, B., Huang, T., 2010. Traffic safety and operations on shared-access facilities: An urban arterial case study. J. Transp. Saf. Secur. 2 (4), 336–351. http://dx.doi.org/10.1080/19439962.2010.522301.

Allen, B.L., Shin, B.T., Cooper, P.J., 1978. Analysis of traffic conflicts and collisions. Transp. Res. Rec. 667 (1), 67–74, URL: https://trid.trb.org/View/85806.

Arun, A., Haque, M.M., Bhaskar, A., Washington, S., Sayed, T., 2021. A systematic mapping review of surrogate safety assessment using traffic conflict techniques. Accid. Anal. Prev. 153, 106016. http://dx.doi.org/10.1016/j.aap.2021.106016.

Bahmankhah, B., Coelho, M.C., 2017. Multi-objective optimization for short distance trips in an urban area: Choosing between motor vehicle or cycling mobility for a safe, smooth and less polluted route. Transp. Res. Procedia 27, 428–435. http://dx.doi.org/10.1016/j.trpro.2017.12.009.

Bahmankhah, B., Fernandes, P., Coelho, M.C., 2019a. Cycling at intersections: a multi-objective assessment for traffic, emissions and safety. Transport 34 (2), 225–236. http://dx.doi.org/10.3846/transport.2019.8946.

Bahmankhah, B., Fernandes, P., Teixeira, J., Coelho, M.C., 2019b. Interaction between motor vehicles and bicycles at two-lane roundabouts: a driving volatility-based analysis. Int. J. Inj. Control. Saf. Promot. 26 (3), 205–215. http://dx.doi.org/10.1080/17457300.2019.1624578.

Ballo, L., de Freitas, L.M., Meister, A., Axhausen, K.W., 2023. The e-bike city as a radical shift toward zero-emission transport: sustainable? equitable? desirable? J. Transp. Geogr. 111, 103663. http://dx.doi.org/10.1016/j.jtrangeo.2023.103663.

Barceló, J., 2010. Models, traffic models, simulation, and traffic simulation. In: Barceló, J. (Ed.), Fundamentals of Traffic Simulation. In: International Series in Operations Research & Management Science, Springer, New York, pp. 1–62. http://dx.doi.org/10.1007/978-1-4419-6142-6_1.

Bjørnskau, T., Aasvik, O., De Ceunynck, T., Fyhri, A., Hagenzieker, M., Johnsson, C., Laureshyn, A., 2023. "Game over" for autonomous shuttles in mixed traffic? Results from field surveys among pedestrians and cyclists on how they interact with autonomous shuttles in real-life traffic in Norway. Transp. Res. Interdiscip. Perspect. 18, http://dx.doi.org/10.1016/j.trip.2023.100781.

Brunner, P., Denk, F., Huber, W., Kates, R., 2019. Virtual safety performance assessment for automated driving in complex urban traffic scenarios. In: 2019 IEEE Intelligent Transportation Systems Conference. ITSC, pp. 679–685. http://dx.doi.org/10.1109/ITSC.2019.8917517.

Buisson, C., Daamen, W., Vincenzo, P., Wagner, P., Montanino, M., Ciuffo, B., 2014. Valibration and validation principles. In: Daamen, W., Buisson, C., Hoogendoorn, S.P. (Eds.), Traffic Simulation and Data: Validation Methods and Applications. CRC Press, Boca Raton, pp. 89–119. http://dx.doi.org/10.1201/b17440.

Campisi, T., Tibljaš, A.D., Tesoriere, G., Canale, A., Rencelj, M., Šurdonja, S., 2020. Cycling traffic at turbo roundabouts: some considerations related to cyclist mobility and safety. Transp. Res. Procedia 45, 627–634. http://dx.doi.org/10.1016/j.trpro.2020.03.048.

COWI, 2012. Mikrosimulering Af Cyclister I Myldretid. Rep. A028928-002, City of Copenhagen.

COWI, 2013. Micro Simulation of Cyclists in Peak Hour Traffic. Rep. A028928-004, City of Copenhagen.

Davis, G.A., Hourdos, J., Xiong, H., Chatterjee, I., 2011. Outline for a causal model of traffic conflicts and crashes. Accid. Anal. Prev. 43 (6), 1907–1919. http://dx.doi.org/10.1016/j.aap.2011.05.001.

Denk, F., Brunner, P., Huber, W., Margreiter, M., Bogenberger, K., Kates, R., 2022. Assessment of traffic safety interventions using virtual randomized controlled trials: Potential of connected and automated driving including V2X for collision reduction at urban intersections. In: 2022 IEEE 25th International Conference on Intelligent Transportation Systems. ITSC, IEEE, pp. 1183–1190. http://dx.doi.org/10.1109/ITSC55140.2022.9921764.

Dijkstra, A., 2012. Effects of a Robust Roads Network on Bicycle Traffic: Results Obtained from a Microsimulation Model. Rep. R-2012-3, SWOV Institute for Road Safety Research, p. 34, URL: https://swov.nl/en/publicatie/effecten-van-een-robuust-wegennet-op-het-fietsverkeer.

Essa, M., Sayed, T., 2015. Transferability of calibrated microsimulation model parameters for safety assessment using simulated conflicts. Accid. Anal. Prev. 84, 41–53. http://dx.doi.org/10.1016/j.aap.2015.08.005.

Fellendorf, M., Vortisch, P., 2011. Microscopic traffic flow simulator VISSIM. In: Barceló, J. (Ed.), Fundamentals of Traffic Simulation. In: International Series in Operations Research & Management Science, Springer, New York, NY, pp. 63–93. http://dx.doi.org/10.1007/978-1-4419-6142-6_2.

Gavriilidou, A., Daamen, W., Yuan, Y., Hoogendoorn, S.P., 2019a. Modelling cyclist queue formation using a two-layer framework for operational cycling behaviour. Transp. Res. Part C: Emerg. Technol. 105, 468–484. http://dx.doi.org/10.1016/j.trc.2019.06.012.

Gavriilidou, A., Yuan, Y., Farah, H., Hoogendoorn, S.P., 2019b. Microscopic cycling behavior model using differential game theory. In: Hamdar, S.H. (Ed.), Traffic and Granular Flow '17. Springer, pp. 497–506. http://dx.doi.org/10.1007/978-3-030-11440-4_54.

German Aerospace Center, 2023a. Bicycles. *SUMO user documentation*. URL: https://sumo.dlr.de/docs/Simulation/Bicycles.html.

German Aerospace Center, 2023b. Safety. *SUMO user documentation*. URL: https://sumo.dlr.de/docs/Simulation/Safety.html.

German Aerospace Center, 2023c. SSM device. *SUMO user documentation*. URL: https://sumo.dlr.de/docs/Simulation/Output/SSM_Device.html.

Gettman, D., Pu, L., Sayed, T., Shelby, S., 2008. Surrogate Safety Assessment Model and Validation: Final Report. Tech. Rep. FHWA-HRT-08-051, Federal Highway Administration, URL: https://www.fhwa.dot.gov/publications/research/safety/08051/index.cfm.

Grigoropoulos, G., Khabibulin, N., Keler, A., Malcolm, P., Bogenberger, K., 2021. Detection and classification of bicyclist group behavior for automated vehicle applications. In: 2021 IEEE International Intelligent Transportation Systems Conference. ITSC, IEEE, pp. 1883–1889. http://dx.doi.org/10.1109/ITSC48978.2021.9564548.

Grigoropoulos, G., Leonhardt, A., Kaths, H., Junghans, M., Baier, M.M., Busch, F., 2022. Traffic flow at signalized intersections with large volumes of bicycle traffic. Transp. Res. Part A: Policy Pr. 155, pp 464–483. http://dx.doi.org/10.1016/j.tra.2021.11.021.

Grigoropoulos, G., Lücken, L., Erdmann, J., Kaths, H., 2019. Modelling bicycle infrastructure in SUMO. In: SUMO User Conference 2019, vol. 62, Berlin, pp. 187–198. http://dx.doi.org/10.29007/6cs5.

Guhathakurta, S., Pande, A., Hwang, U., Lee, K., Molan, A., Sergio, B., 2023. Expanding Mobility Options for All: Optimizing and Extending the Biking Infrastructure to Generate Complete Street Networks in Atlanta. Tech. Rep. CTEDD 021-04, Center for Transportation, Equity, Decisions and Dollars (CTEDD), URL: https://rosap.ntl.bts.gov/view/dot/72382.

Guo, Y., Essa, M., Sayed, T., Haque, M.M., Washington, S., 2019. A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in Australia. Transp. Res. Part C: Emerg. Technol. 101, 96–110. http://dx.doi.org/10.1016/j.trc.2019.02.009.

Hayward, J., 1972. Near-miss determination through use of a scale of danger. In: 51st Annual Meeting of the Highway Research Board. In: Highway Research Record, vol. 384, pp. 24–34, URL: http://onlinepubs.trb.org/Onlinepubs/hrr/1972/384/384-004.pdf.

Hollander, Y., Liu, R., 2008. The principles of calibrating traffic microsimulation models. Transportation 35 (3), 347–362. http://dx.doi.org/10.1007/s11116-007-9156-2.

Hoogendoorn, S., Gavriilidou, A., Daamen, W., Duives, D., 2021. Game theoretical framework for bicycle operations: A multi-strategy framework. Transp. Res. Part C: Emerg. Technol. 128, 103175. http://dx.doi.org/10.1016/j.trc.2021.103175.

Huang, F., Liu, P., Yu, H., Wang, W., 2013. Identifying if VISSIM simulation model and SSAM provide reasonable estimates for field measured traffic conflicts at signalized intersections. Accid. Anal. Prev. 50, 1014–1024. http://dx.doi.org/10.1016/j.aap.2012.08.018.

Hübner, K., Schünemann, B., Schilling, T., Radusch, I., 2017. On assessing road safety aspects of a cycling router application. In: 2017 15th International Conference on ITS Telecommunications. ITST, IEEE, Warsaw, Poland, pp. 1–8. http://dx.doi.org/10.1109/ITST.2017.7972199.

Hydén, C., 1987. The development of a method for traffic safety evaluation: The Swedish traffic conflicts technique. In: Bulletin Lund Institute of Technology, Department of Traffic Planning and Engineering, vol. 70, pp. 1–229.

Janssen, C.P., Boyle, L.N., Ju, W., Riener, A., Alvarez, I., 2020. Agents, environments, scenarios: A framework for examining models and simulations of human-vehicle interaction. Transp. Res. Interdiscip. Perspect. 8, 100214. http://dx.doi.org/10.1016/j.trip.2020.100214.

Jiang, J., Ren, Y., Guan, Y., Eben Li, S., Yin, Y., Yu, D., Jin, X., 2022. Integrated decision and control at multi-lane intersections with mixed traffic flow. J. Phys.: Conf. Ser. 2234 (1), 012015. http://dx.doi.org/10.1088/1742-6596/2234/1/012015.

Johnsson, C., Laureshyn, A., Dágostino, C., 2021. Validation of surrogate measures of safety with a focus on bicyclist–motor vehicle interactions. Accid. Anal. Prev. 153, 106037. http://dx.doi.org/10.1016/j.aap.2021.106037.

Johnsson, C., Laureshyn, A., De Ceunynck, T., 2018. In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators. Transp. Rev. 38 (6), 765–785. http://dx.doi.org/10.1080/01441647.2018.1442888.

Joo, S.H., Oh, C., Lee, S.S., 2012. Evaluation of left-turn passages for bicycle traffic in mixed traffic stream at signalized intersections. Int. J. Highw. Eng. 14 (5), 145–155. http://dx.doi.org/10.7855/IJHE.2012.14.5.145.

Karkhanis, P., Dajsuren, Y., van den Brand, M.G., Josten, J., 2020. Modelling and simulating safety and efficiency at an unsignalised intersection. In: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems. ITSC, pp. 1–6. http://dx.doi.org/10.1109/ITSC45102.2020.9294193.

Kaths, H., 2023. A movement and interaction model for cyclists and other non-lane-based road users. Front. Futur. Transp. 4, http://dx.doi.org/10.3389/ffutr.2023.1183270.

Kaths, H., Keler, A., Bogenberger, K., 2021. Calibrating the wiedemann 99 car-following model for bicycle traffic. Sustainability 13 (6), 3487. http://dx.doi.org/10.3390/su13063487.

Kesting, A., Treiber, M., Helbing, D., 2010. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. 368, 4585–4605. http://dx.doi.org/10.1098/rsta.2010.0084.

Kodupuganti, S.R., Pulugurtha, S.S., 2022. Modeling safety performance of urban roads with heterogeneous traffic conditions. In: ASCE International Conference on Transportation and Development. ASCE, Seattle, pp. 197–208. http://dx.doi.org/10.1061/9780784484333.018.

Krauß, S., 1998. Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics. (Doctoral dissertation). Universität zu Köln, Köln, Germany, URL: https://sumo.dlr.de/pdf/KraussDiss.pdf.

Langer, M., Kates, R., Bogenberger, K., 2023. Simulation of urban crash occurrence based on real-world crash data. Transp. Res. Rec. 2677 (2), 1150–1164. http://dx.doi.org/10.1177/03611981221112400.

Laureshyn, A., de Goede, M., Saunier, N., Fyhri, A., 2017. Cross-comparison of three surrogate safety methods to diagnose cyclist safety problems at intersections in Norway. Accid. Anal. Prev. 105, 11–20. http://dx.doi.org/10.1016/j.aap.2016.04.035.

Ledezma-Navarro, B., Stipancic, J., Andreoli, A., Miranda-Moreno, L., 2018. Evaluation of level of service and safety for vehicles and cyclists at signalized intersections. In: Transportation Research Board 97th Annual Meeting. Washington D.C., URL: https://trid.trb.org/view/1496416.

Lemcke, D., Riffle, K., Russo, B.J., Smaglik, E.J., 2021. Examining the use of microsimulation modeling to assess bicycle-vehicle conflicts at intersections: A case study incorporating field-observed conflict data. In: Transportation Research Board 100th Annual Meeting. TRB, Washington D.C., URL: https://trid.trb.org/view/1759925.

Li, Y., Ni, Y., Sun, J., Ma, Z., 2020. Modeling the illegal lane-changing behavior of bicycles on road segments: Considering lane-changing categories and bicycle heterogeneity. Phys. A 541, 123302. http://dx.doi.org/10.1016/j.physa.2019.123302.

Li, M., Shi, F., Chen, D., 2011. Analyze bicycle-car mixed flow by social force model for collision risk evaluation. In: 3rd International Conference on Road Safety and Simulation. TRB, Indianapolis, URL: https://trid.trb.org/view/1285680.

van Lint, J.W.C., Calvert, S.C., 2018. A generic multi-level framework for microscopic traffic simulation—Theory and an example case in modelling driver distraction. Transp. Res. Part B: Methodol. 117, 63–86. http://dx.doi.org/10.1016/j.trb.2018.08.009.

Liu, Q., Sun, J., Tian, Y., Xiong, L., 2020. Modeling and simulation of overtaking events by heterogeneous non-motorized vehicles on shared roadway segments. Simul. Model. Pr. Theory 103, 102072. http://dx.doi.org/10.1016/j.simpat.2020.102072.

Lopez, P.A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., Wiessner, E., 2018. Microscopic traffic simulation using SUMO. In: 2018 21st International Conference on Intelligent Transportation Systems. ITSC, IEEE, Maui, USA, pp. 2575–2582. http://dx.doi.org/10.1109/ITSC.2018.8569938.

Lu, F., He, Z., 2019. Case study on traffic organization optimization in the roads around a middle school. In: CICTP 2019. ASCE, Nanjing, China, pp. 2543–2555. http://dx.doi.org/10.1061/9780784482292.221.

Ma, Z., Sun, J., Wang, Y., 2017. A two-dimensional simulation model for modelling turning vehicles at mixed-flow intersections. Transp. Res. Part C: Emerg. Technol. 75, 103–119. http://dx.doi.org/10.1016/j.trc.2016.12.005.

Monsere, C., McNeil, N., Wang, Y., Sanders, R., Burchfield, R., Schultheiss, W., 2019. Contextual Guidance at Intersections for Protected Bicycle Lanes. Tech. Rep. NITC-RR-987, Portland State University, Transportation Research and Education Center (TREC), http://dx.doi.org/10.15760/trec.241.

Nabavi Niaki, M.S., Saunier, N., Miranda-Moreno, L.F., 2018. Analysing cyclist behaviour at cycling facility discontinuities using video data. Trans. Transp. Sci. 9 (1), 3–17. http://dx.doi.org/10.5507/tots.2018.006.

Nguyen, J., Powers, S.T., Urquhart, N., Farrenkopf, T., Guckert, M., 2021. An overview of agent-based traffic simulators. Transp. Res. Interdiscip. Perspect. 12, 100486. http://dx.doi.org/10.1016/j.trip.2021.100486.

Ni, Y., Li, Y., Yuan, Y., Sun, J., 2023. An operational simulation framework for modelling the multi-interaction of two-wheelers on mixed-traffic road segments. Phys. A 611, http://dx.doi.org/10.1016/j.physa.2022.128441.

Ni, Y., Xin, L., Long, K., 2019. Modeling and simulation of right-turning vehicle-bicycle interactions at intersections. J. Tongij Univ. (Natural Science) 47 (7), 0994–1003. http://dx.doi.org/10.11908/j.issn.0253-374x.2019.07.011.

Pauwels, A., Pourmohammad-Zia, N., Schulte, F., 2022. Safety and sustainable development of automated driving in mixed-traffic urban areas—Considering vulnerable road users and network efficiency. Sustainability 14 (20), 13486. http://dx.doi.org/10.3390/su142013486.

Pechinger, M., Schroer, G., Bogenberger, K., Markgraf, C., 2021. Cyclist safety in urban automated driving - sub-microscopic HIL simulation. In: 2021 IEEE International Intelligent Transportation Systems Conference. ITSC, IEEE, Indianapolis, pp. 615–620. http://dx.doi.org/10.1109/ITSC48978.2021.9565108.

Pechinger, M., Schröer, G., Bogenberger, K., Markgraf, C., 2023. Roadside infrastructure support for Urban Automated Driving. IEEE Trans. Intell. Transp. Syst. 24 (10), 10643–10652. http://dx.doi.org/10.1109/TITS.2023.3277138.

Preston, A., Pulugurtha, S.S., 2021. Simulating and assessing the effect of a protected intersection design for bicyclists on traffic operational performance and safety. Transp. Res. Interdiscip. Perspect. 9 (1), 100329. http://dx.doi.org/10.1016/j.trip.2021.100329.

PTV Group, 2023. PTV vissim 2023. Software Manual 2023.00-01. PTV Planung Transport Verkehr GmbH.

Pucher, J., Buehler, R., 2021. Introduction: Cycling to sustainability. In: Buehler, R., Pucher, J., Gottlieb, R., Luce, H.R. (Eds.), Cycling for Sustainable Cities. In: Urban and Industrial Environments, MIT Press, pp. 1–10. http://dx.doi.org/10.7551/mitpress/11963.003.0005.

Qian, X., Jaller, M., Xiao, I.R., Chen, S., 2022. Analysis of Intelligent Vehicle Technologies to Improve Vulnerable Road Users Safety at Signalized Intersections. Tech. Rep. NCST-UCD-RR-22-28, UC Davis, National Center for Sustainable Transportation (NCST), p. 83, URL: https://rosap.ntl.bts.gov/view/dot/63027.

Ren, G., Jiang, H., Chen, J., Huang, Z., Lu, L., 2016. Heterogeneous cellular automata model for straight-through bicycle traffic at signalized intersection. Phys. A 451, 70–83. http://dx.doi.org/10.1016/j.physa.2015.12.159.

Ren, Y., Jiang, J., Zhan, G., Li, S.E., Chen, C., Li, K., Duan, J., 2022. Self-learned intelligence for integrated decision and control of automated vehicles at signalized intersections. IEEE Trans. Intell. Transp. Syst. 23 (12), 24145–24156. http://dx.doi.org/10.1109/TITS.2022.3196167.

Ren, Y., Zhan, G., Tang, L., Li, S.E., Jiang, J., Li, K., Duan, J., 2023. Improve generalization of driving policy at signalized intersections with adversarial learning. Transp. Res. Part C: Emerg. Technol. 152 (1), 104161. http://dx.doi.org/10.1016/j.trc.2023.104161.

Rinke, N., Schiermeyer, C., Pascucci, F., Berkhahn, V., Friedrich, B., 2017. A multi-layer social force approach to model interactions in shared spaces using collision prediction. Transp. Res. Procedia 25 (1), 1249–1267. http://dx.doi.org/10.1016/j.trpro.2017.05.144.

Roosta, A., Kaths, H., Barthauer, M., Erdmann, J., Flötteröd, Y.P., Behrisch, M., 2023. The state of bicycle modeling in SUMO. SUMO Conf. Proc. 4, 55–64. http://dx.doi.org/10.52825/scp.v4i.215.

Russo, B., Hurwitz, D.S., Kothuri, S., Smaglik, E., Scott-Deeter, L., Lemcke, D., Yates, E., 2022. Impacts of Intersection Treatments and Traffic Characteristics on Bicyclist Safety. Tech. Rep. FHWA-OR-RD-23-02, Oregon Department of Transportation, Research Section, URL: https://rosap.ntl.bts.gov/view/dot/63551.

Sargent, R.G., 2010. Verification and validation of simulation models. In: Proceedings of the 2010 Winter Simulation Conference. IEEE, Baltimore (MD), USA, pp. 166–183. http://dx.doi.org/10.1109/WSC.2010.5679166.

Schmidt, C.M., Dabiri, A., Schulte, F., Happee, R., Moore, J., 2024. Essential bicycle dynamics for microscopic traffic simulation: An example using the social force model. In: The Evolving Scholar - BMD 2023, fifth ed. TU Delft Open, http://dx.doi.org/10.59490/65a5124da90ad4aecf0ab147.

Semrau, M., Erdmann, J., Friedrich, B., Waldmann, R., 2016. Simulation framework for testing ADAS in Chinese traffic situations. In: SUMO 2016 – Traffic, Mobility, and Logistics, vol. 30, Berlin, pp. 103–114, URL: https://elib.dlr.de/104432/.

Siebke, C., Bäumler, M., Blenz, K., Lehmann, M., Ringhand, M., Mai, M., Prokop, G., 2023. Predicting the impact on road safety of an intersection AEB at urban intersections. using a novel virtual test field for the assessment of conflict prevention between cyclists/pedelecs and cars. Transp. Res. Interdiscip. Perspect. 17, 100728. http://dx.doi.org/10.1016/j.trip.2022.100728.

Silva, J., Barua, V., Silvera, M., Campos, F., 2023. Analysis of conflicts and delays of cyclists, vehicles, and pedestrians at a busy intersection in Lima Downtown, Peru. In: 2023 Congreso Internacional de Innovación Y Tendencias En IngenierÍa. CONIITI, pp. 1–6. http://dx.doi.org/10.1109/CONIITI61170.2023.10324228.

Sohrabi, S., Khodadadi, A., Mousavi, S., Dadashova, B., Lord, D., 2021. Quantifying the automated vehicle safety performance: A scoping review of the literature, evaluation of methods, and directions for future research. Accid. Anal. Prev. 152, http://dx.doi.org/10.1016/j.aap.2021.106003.

Sun, J., Liu, H., Ma, Z., 2019. Modelling and simulation of highly mixed traffic flow on two-lane two-way urban streets. Simul. Model. Pr. Theory 95, 16–35. http://dx.doi.org/10.1016/j.simpat.2019.04.005.

Sun, J., Qi, X., Xu, Y., Tian, Y., 2020. Vehicle turning behavior modeling at conflicting areas of mixed-flow intersections based on deep learning. IEEE Trans. Intell. Transp. Syst. 21 (9), 3674–3685. http://dx.doi.org/10.1109/TITS.2019.2931701.

Tafidis, P., Pirdavani, A., Brijs, T., Farah, H., 2019. Can automated vehicles improve cyclist safety in urban areas? Safety 5 (3), 57. http://dx.doi.org/10.3390/safety5030057.

Tarko, A., 2018. Surrogate measures of safety. In: Lord, D., Washington, S. (Eds.), Safe Mobility: Challenges, Methodology and Solutions. In: Transport and Sustainability, vol. 11, Emerald Publishing Limited, Bingley, pp. 383–405. http://dx.doi.org/10.1108/S2044-994120180000011019.

Thompson, J., Read, G.J.M., Wijnands, J.S., Salmon, P.M., 2020. The perils of perfect performance; considering the effects of introducing autonomous vehicles on rates of car vs cyclist conflict. Ergonomics 63 (8), 981–996. http://dx.doi.org/10.1080/00140139.2020.1739326.

Thompson, J., Savino, G., Stevenson, M., 2015. Reconsidering the safety in numbers effect for vulnerable road users: An application of agent-based modeling. Traffic Inj. Prev. 16 (2), 147–153. http://dx.doi.org/10.1080/15389588.2014.914626.

Thompson, J., Savino, G., Stevenson, M., 2016. A model of behavioural adaptation as a contributor to the safety-in-numbers effect for cyclists. Transp. Res. Part A: Policy Pr. 85, 65–75. http://dx.doi.org/10.1016/j.tra.2015.12.004.

Thompson, J., Wijnands, J.S., Savino, G., Lawrence, B., Stevenson, M., 2017. Estimating the safety benefit of separated cycling infrastructure adjusted for behavioral adaptation among drivers; an application of agent-based modelling. Transp. Res. Part F: Traffic Psychol. Behav. 49, 18–28. http://dx.doi.org/10.1016/j.trf.2017.05.006.

Twaddle, H., Busch, F., 2019. Binomial and multinomial regression models for predicting the tactical choices of bicyclists at signalised intersections. Transp. Res. Part F: Traffic Psychol. Behav. 60, 47–57. http://dx.doi.org/10.1016/j.trf.2018.10.002.

Twaddle, H., Schendzielorz, T., Fakler, O., 2014. Bicycles in urban areas: Review of existing methods for modeling behavior. Transp. Res. Rec. 2434 (1), 140–146. http://dx.doi.org/10.3141/2434-17.

Vanparijs, J., Int Panis, L., Meeusen, R., de Geus, B., 2015. Exposure measurement in bicycle safety analysis: A review of the literature. Accid. Anal. Prev. 84, 9–19. http://dx.doi.org/10.1016/j.aap.2015.08.007.

Wallentin, G., Loidl, M., 2016. Bicycle-bicycle accidents emerge from encounters: An agent-based approach. Safety 2 (2), 14. http://dx.doi.org/10.3390/safety2020014.

Wiedemann, R., 1974. Simulation des strassenverkehrsflusses. In: Schriftenreihe des Instituts für Verkehrswesen der Universität Karlsruhe. (Doctoral dissertation). Universität Karlsruhe, Karlsruhe Institute of Technology.

Wu, J., Radwan, E., Abou-Senna, H., 2018. Determination if VISSIM and SSAM could estimate pedestrian-vehicle conflicts at signalized intersections. J. Transp. Saf. Secur. 10 (6), 572–585. http://dx.doi.org/10.1080/19439962.2017.1333181.

Wu, Z., Zhao, L., Sun, J., 2014. A study of dynamic right-turn signal control strategy at mixed traffic flow intersections. Promet – Traffic Transp. 26, 449–458. http://dx.doi.org/10.7307/ptt.v26i6.1357.

Wunderlich, K.E., Vasudevan, M., Wang, P., Dowling, R., Skabardonis, A., Alexiadis, V., 2019. Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software - 2019 Update to the 2004 Version. Tech. Rep. FHWA-HOP-18-036, FHWA, URL: https://ops.fhwa.dot.gov/publications/fhwahop18036/index.htm.

Xu, Z., Zheng, N., Logan, D., Vu, H., 2023. Assessing bicycle-vehicle conflicts at urban intersections utilizing a VR integrated simulation approach. Accid. Anal. Prev. 191, http://dx.doi.org/10.1016/j.aap.2023.107194.

Young, W., Sobhani, A., Lenné, M.G., Sarvi, M., 2014. Simulation of safety: A review of the state of the art in road safety simulation modelling. Accid. Anal. Prev. 66, 89–103. http://dx.doi.org/10.1016/j.aap.2014.01.008.

Zhou, W., Wang, X., 2022. Calibrating and comparing autonomous braking systems in motorized-to-non-motorized-vehicle conflict scenarios. IEEE Trans. Intell. Transp. Syst. 23 (11), 20636–20651. http://dx.doi.org/10.1109/TITS.2022.3170978.