Department of Transportation Planning and Engineering
School of Civil Engineering
National Technical University of Athens

Thesis submitted for the degree of
Doctor of Philosophy

22 January 2026

Multilevel crash risk assessment under autonomous vehicle integration: From partial to full autonomy through urban traffic simulation

Maria G. Oikonomou

Advisory Committee:

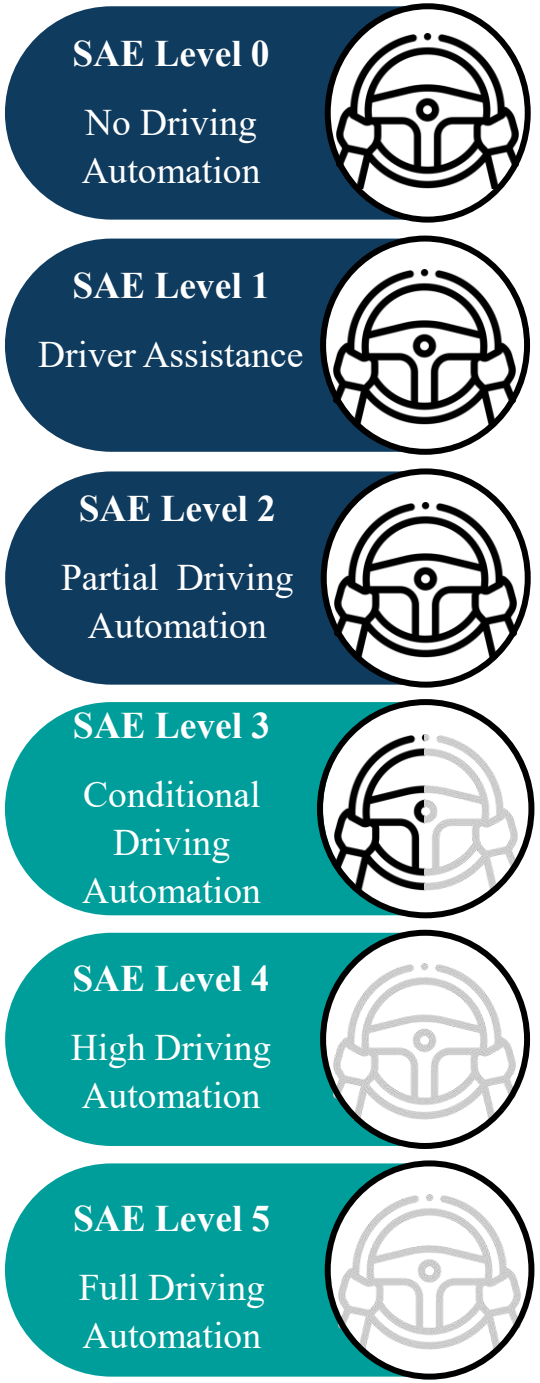
Prof. George Yannis (Supervisor)
Prof. Eleni Vlahogianni
Asst. Prof. Konstantinos Gkiotsalitis

Overview

1. Background and research motivation
2. Literature review and behavioural modelling
3. Methodological framework and simulation setup
4. Crash risk estimation and validation
5. Multilevel safety assessment and key findings
6. Implications, limitations and future research

Introduction

- Road crashes cause over **1 million fatalities annually** (1.19 M in 2023), with human error responsible for 94% of crashes.
- Autonomous Vehicles (AVs) are expected to **reduce human error** while improving traffic safety and mobility.
- Driving automation is classified into **SAE Levels 0-5**, which distinguish vehicles based on how driving responsibilities are shared between the human driver and the automated system.
- Current real-world deployment is dominated by **SAE Level 2**, with limited Level 3 pilots, Level 4 operating only in geo-fenced or controlled environments and Level 5 remaining a long-term objective.



Role of simulation in safety assessment

Why simulation?

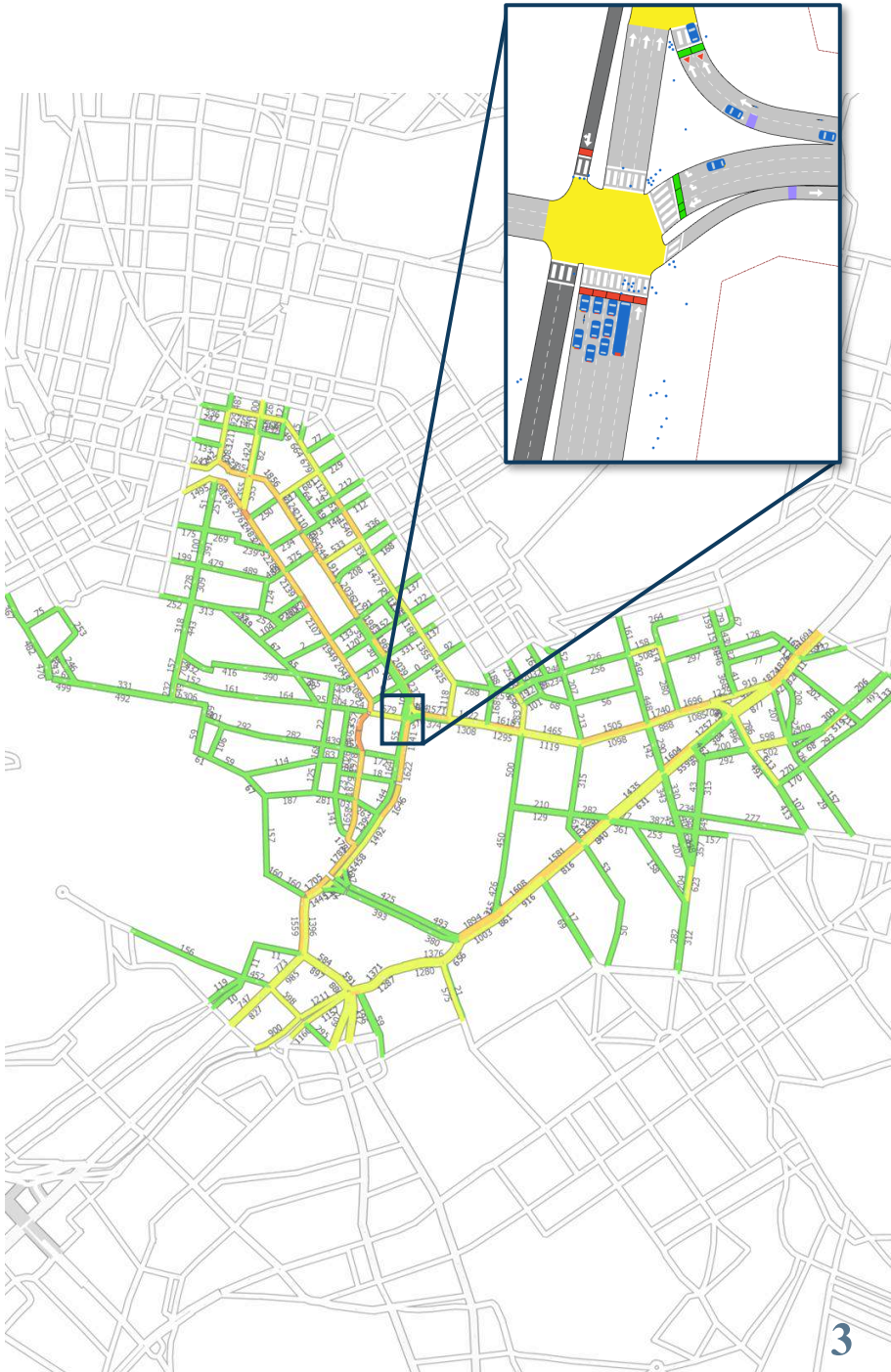
- Lack of historical crash data for high level AVs and high market penetration scenarios.
- Real-world testing is limited geographically.

Microscopic simulation advantages

- Allows testing of complex and high-risk scenarios.
- Enables evaluation of design factors.

Surrogate Safety Measures (SSMs)

- Quantify safety without needing crash data.
- Strong conceptual link to crash likelihood and severity.



Problem statement

1. Ensuring AV safety remains challenging despite their potential to reduce human error, evaluation must cover diverse real-world scenarios **across all SAE levels**.

2. Lack of historical crash data for higher level AVs forces reliance on microscopic simulation and surrogate safety measures, yet **crash-free simulations limit realism**.

3. Uncertainty in choosing **appropriate simulation parameters** due to limited research and inconsistent modelling approaches across studies and software tools.

4. Simulation limitations include difficulty in replicating human behaviour, complex environments and unpredictable events, making **real-world validation** essential.

Objectives

01

Review of existing literature to **define parameters utilized in modelling AVs** within traffic simulation.

02

Identify the **applications of traffic simulation on road safety assessment** by synthesizing past research findings.

03

Establish a **solid microsimulation-based approach** for estimating crash risk.

04

Validate this crash-risk evaluation method by comparing simulated safety metrics with field crash data.

05

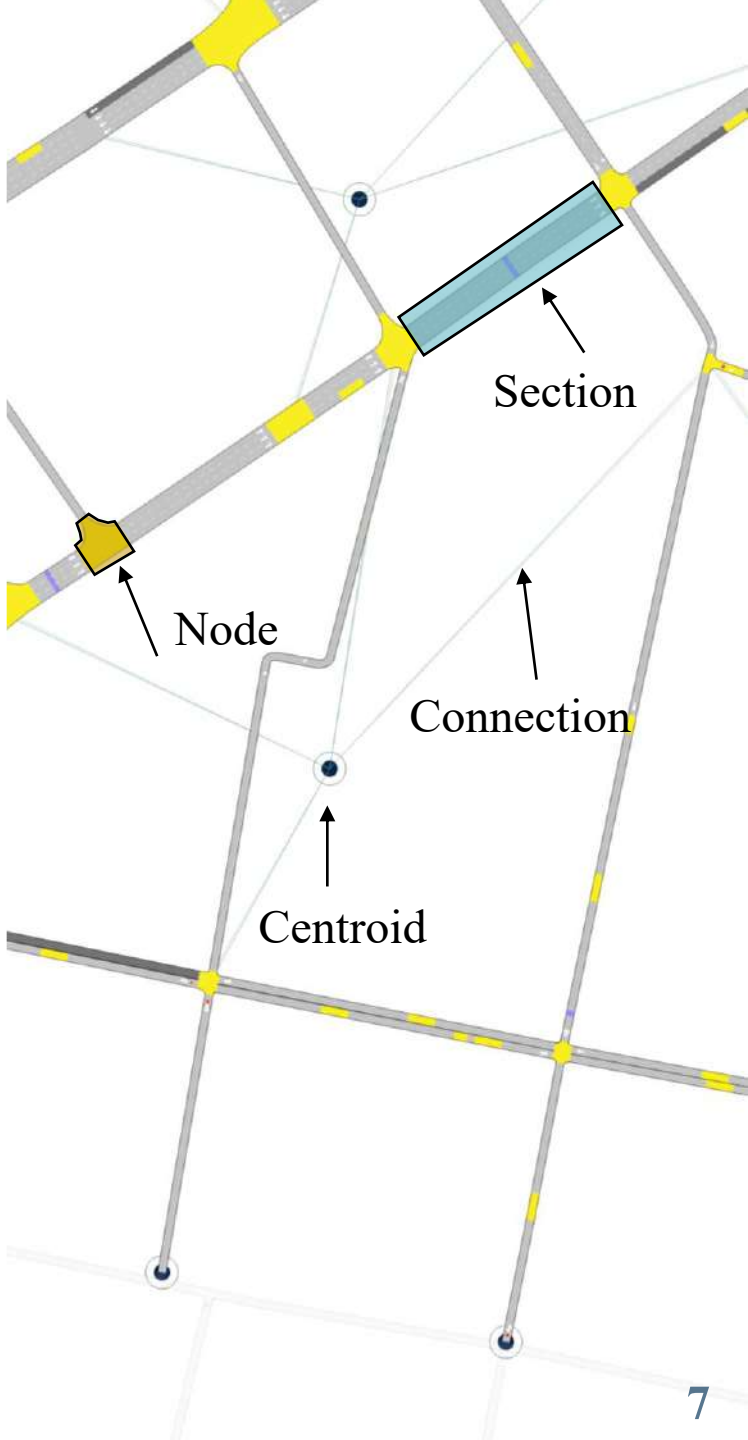
Conduct a **multilevel safety assessment of AV deployment** using an integrated simulation-Machine Learning (ML)-spatial framework.

Research questions

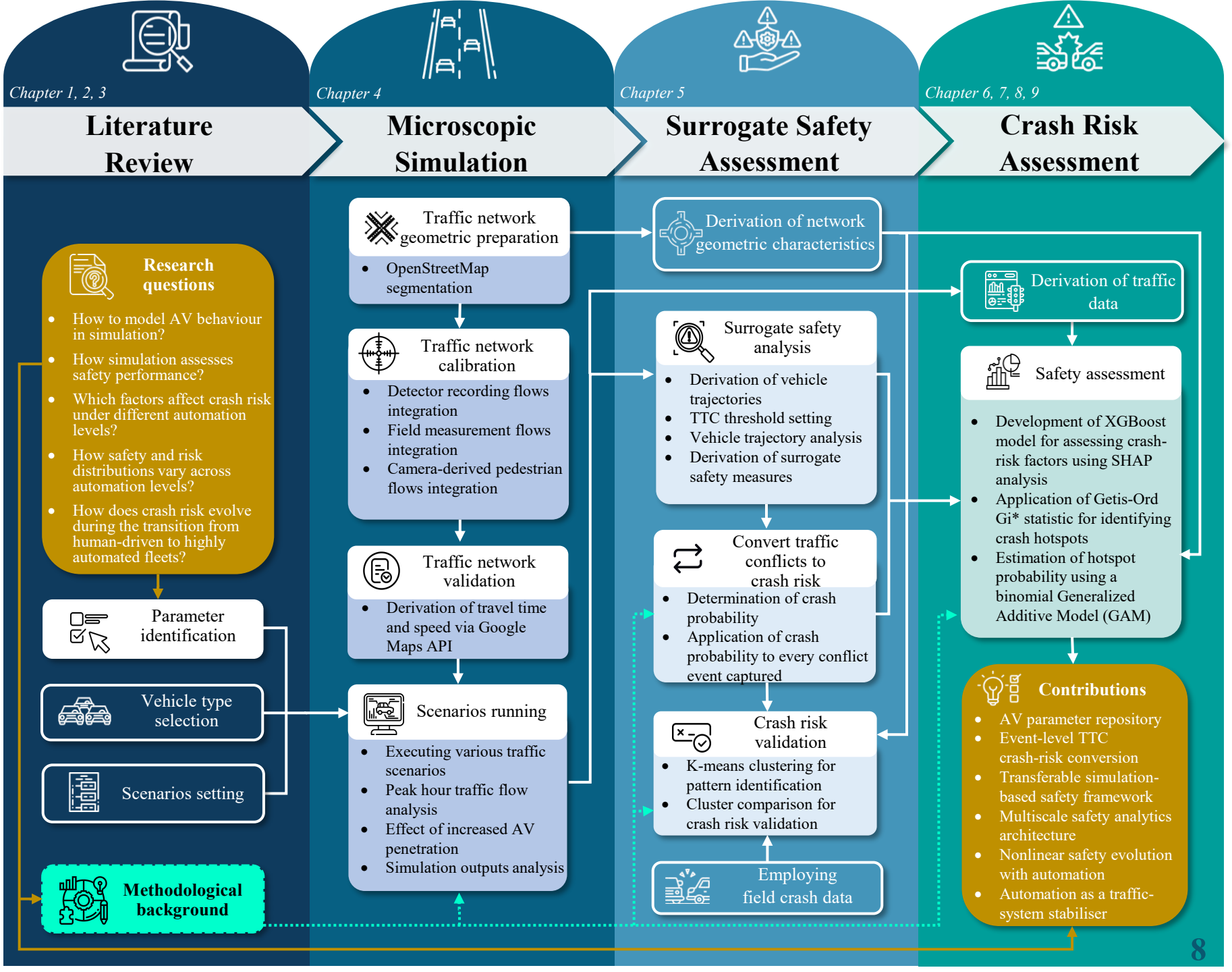
RQ1	RQ2	RQ3	RQ4	RQ5
Which **parameters** can represent and **model the behaviour** of autonomous vehicles within traffic microsimulation frameworks?	How can traffic **microsimulation** be effectively used to **assess road safety** performance?	Which design, traffic, and automation-related **factors most critically affect road safety** outcomes across different automation levels and market penetration rates in urban traffic systems?	How do **safety patterns and risk distributions change** across different fleets of automation, and at multiple scales within urban road networks?	How does **crash risk evolve** during the transition from human-driven traffic to highly automated fleets?

Basics of microscopic simulation

- Microscopic simulation represents traffic at the level of **individual vehicles** and operates in discrete time steps.
- At each **simulation step**, vehicle speed, acceleration and position are updated based on vehicle behaviour.
- Vehicle movement is governed by **behavioral models** acting on network elements such as sections and nodes.
- In this study, demand is defined using **Origin-Destination (OD) matrices** and trips are generated and attracted at centroids.

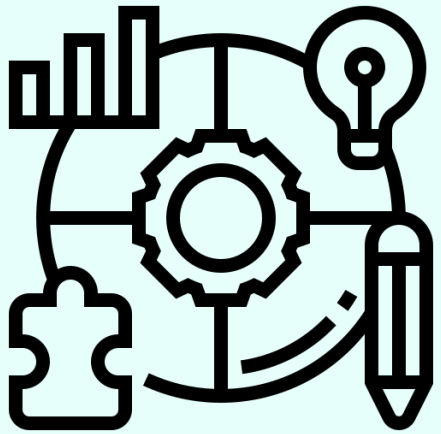


Methodological framework



Methodological background

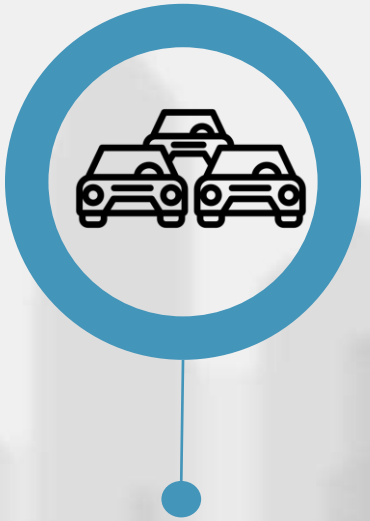
- Prediction used **statistical and ML models**, including linear and logistic regression, spatial logistic regression, Generalized Additive Model (GAM) for nonlinear effects and tree-based models (Random Forest and XGBoost), with SHAP for interpretability.
- K-means **clustering** (with K selected via the Elbow method) was used for pattern identification, with Linear Discriminant Analysis (LDA) for cluster discrimination.
- Model **performance** was evaluated using regression (MAE, RMSE, MSE, R^2/Adjusted R^2) and classification (confusion matrix, precision, recall, F1Score, AUC/ROC, Precision-Recall/Average Precision) metrics.
- **Spatial** diagnostics were conducted using the Getis-Ord Gi* statistic to identify crash risk hotspots.
- Simulation model **calibration** used the GEH statistic to compare simulated and observed flows.



Literature review

Literature review of parameter selection

Parameter identification

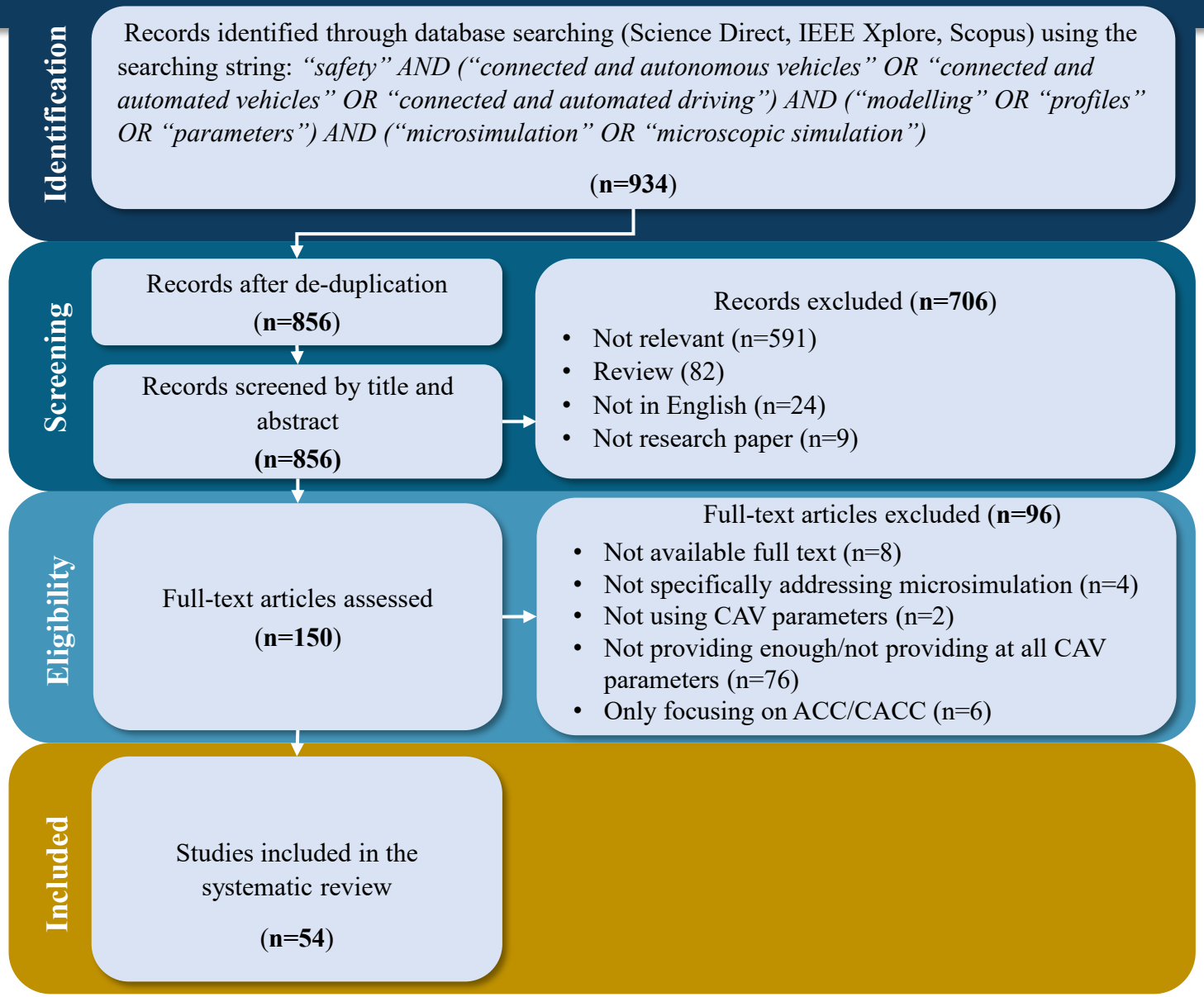


Vehicle type selection

Scenarios setting

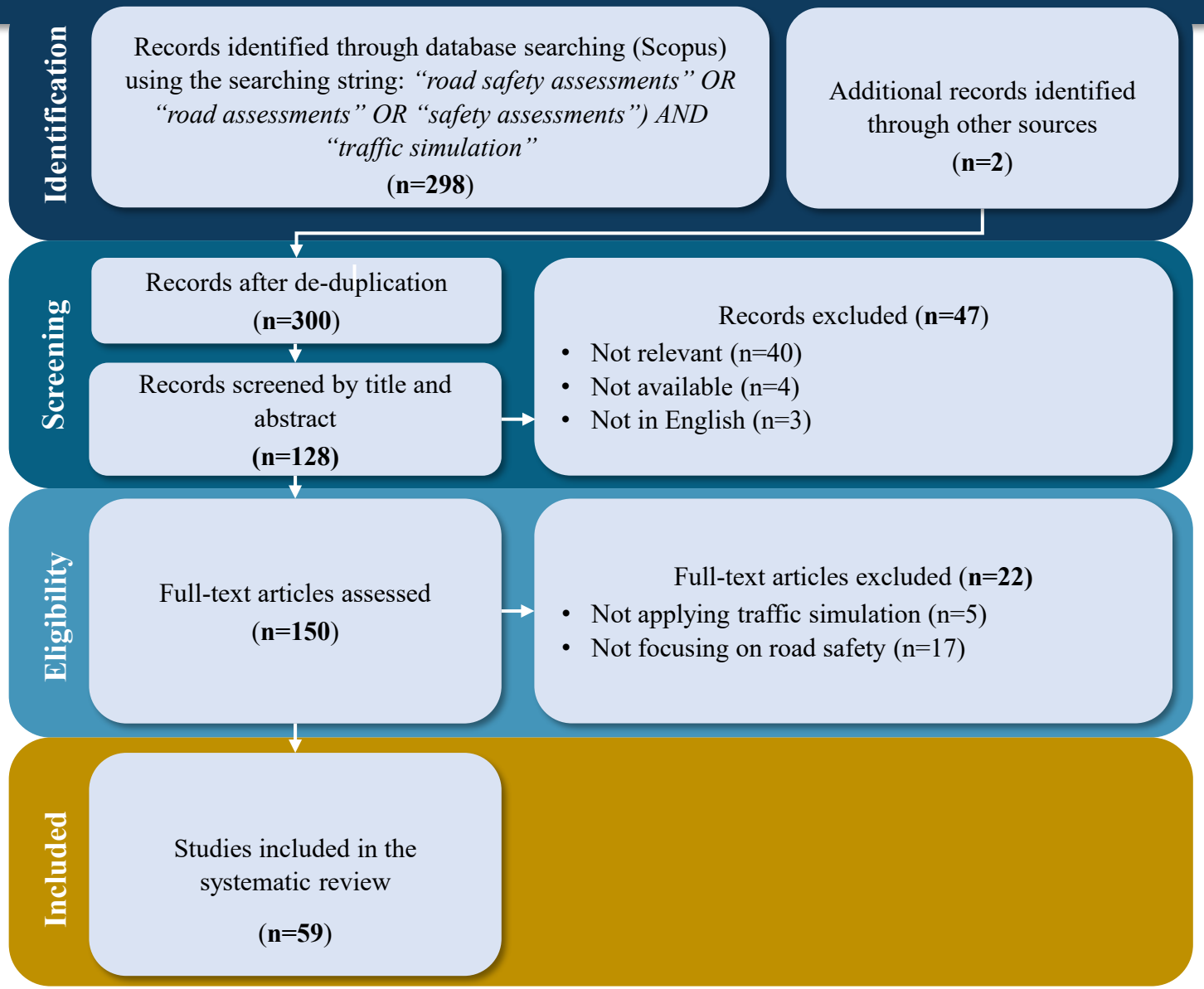
Literature review (1/2)

- PRISMA-based review screened 934 studies and identified 54 studies providing **AV behavioural parameters**.
- Parameters **vary** by SAE level, driving style (e.g. cautious/aggressive) and connectivity.
- Most studies adapt **human-driven vehicle-based models** (e.g. Wiedemann, IDM, Gipps) using adjusted AV parameters.
- Parameter **sources** commonly based on theoretical assumptions, transferred defaults, or small-scale experiments.



Literature review (2/2)

- PRISMA-based review identified extensive **use of microsimulation** for road safety assessment.
- **Conflict-based approaches** dominate the literature, enabling safety analysis across intersections, urban networks, highways and emerging technologies, including automated mobility.
- **Automation-focused studies** remain heterogeneous, with wide variation in behavioural assumptions and parameter choices.



Vehicle type selection

- Literature showed many AV types (different SAE levels, connectivity, aggressiveness, etc.).
- This thesis selects representative vehicle types covering future technologies (SAE-based groups):
  - **SAE Level 0-1**: No or minimal automation; full driver control (serves as the baseline HDV reference).
  - **SAE Level 2-3**: Systems assist with tasks such as ACC and lane keeping and driver must stay engaged.
  - **SAE Level 4-5**: High to complete autonomy under defined or all conditions.
  - **SAE Level 4-5 (aggressive mode)**: Fully automated vehicles with aggressive performance settings.

Parameter identification

RQ1

Car following

- Depends on parameters such as driver characteristics and road geometry
- Comprises acceleration and deceleration components, involving speed adjustments influenced by surrounding conditions

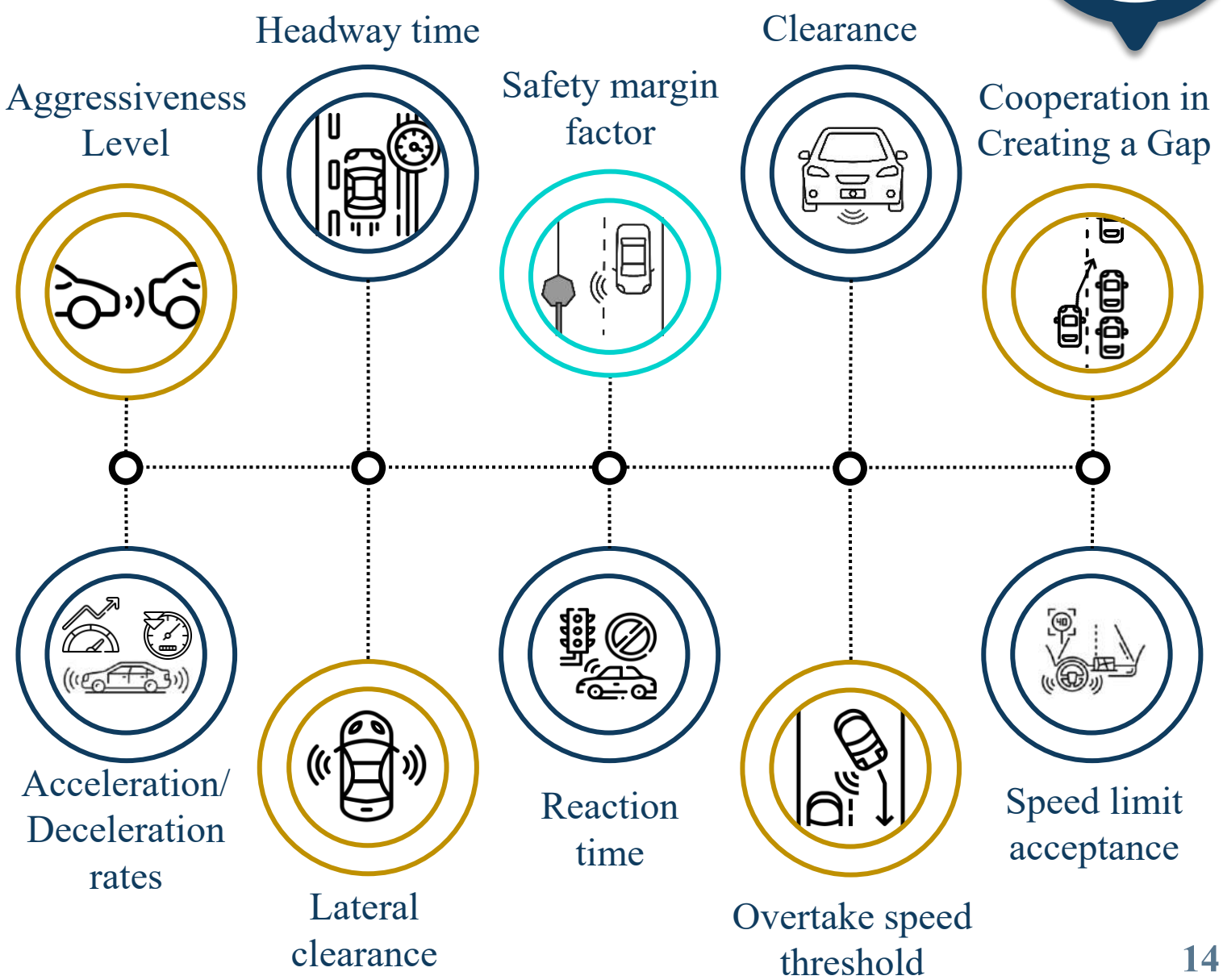
Lane changing

A lane change is made depending on:

- Need to change lane:
  - Target lanes
  - Distance to obstacle
  - Overtaking
- Desirability of lane change
- Feasibility of lane change

Gap acceptance

- Governs the decision to cross
- Applied to Yield and Stop signs
- Assesses intersection crossing based on nearby vehicles



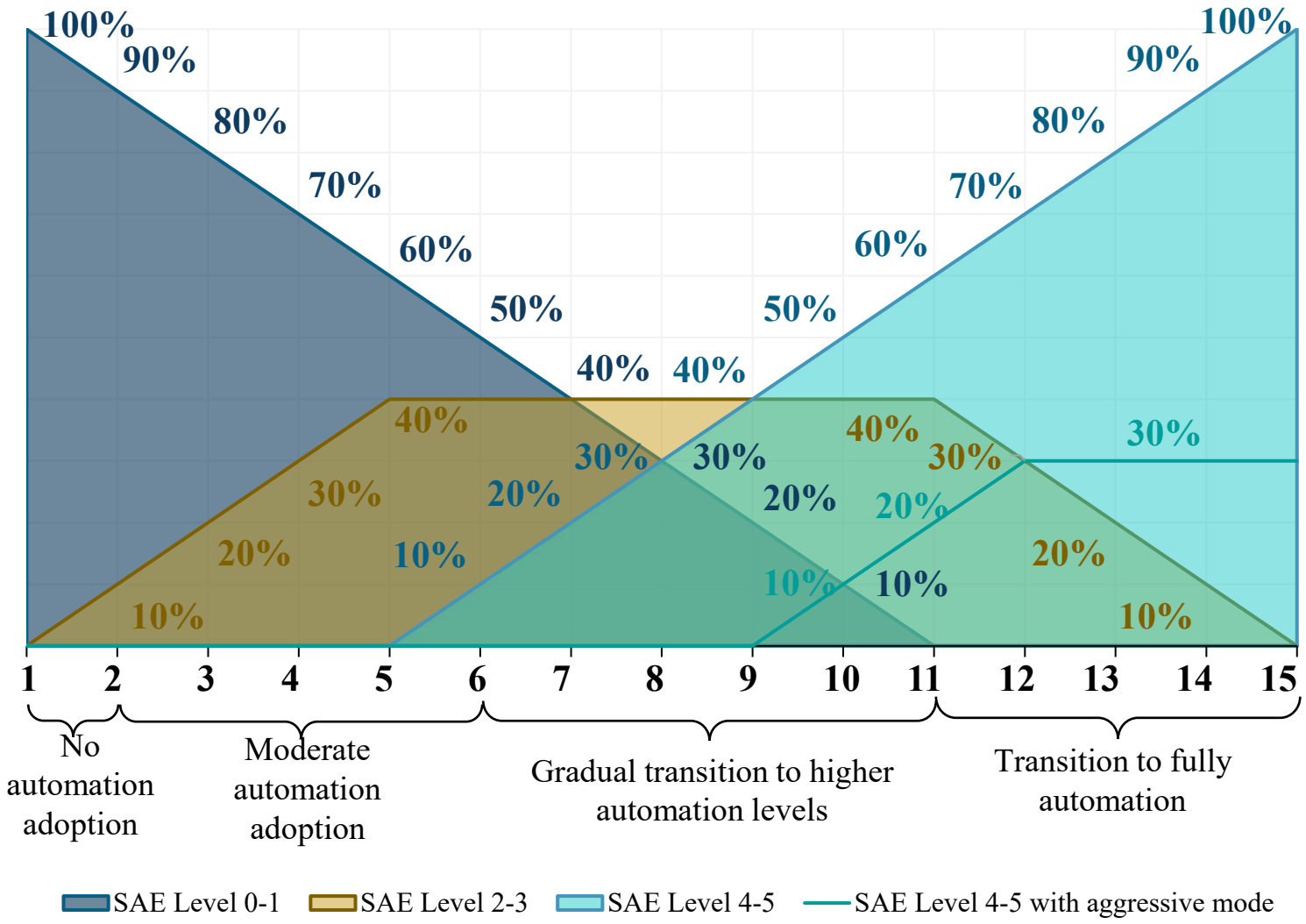
Parameter values

- For several parameters, **SAE 2-3 deviates more** from SAE 0-1 than standard SAE 4-5 (e.g., shorter headways, stronger braking, lower safety margins), reflecting assisted driving without conservative safety policies, whereas SAE 4-5 is modelled with comfort and mixed-traffic safety constraints; the aggressive SAE 4-5 variant isolates assertive behaviour.

Parameter		Vehicle type	Mean	Sd	Min	Max
Maximum acceleration (m/s²)	The maximum rate at which a vehicle can increase its velocity.	SAE level 0-1	2.27	1.33	0.40	5.00
		SAE level 2-3	2.11	0.58	1.40	3.50
		SAE level 4-5	2.23	1.16	0.40	6.00
		SAE level 4-5 (aggressive mode)	2.92	0.93	2.00	4.00
Maximum deceleration (m/s²)	The maximum rate at which a vehicle can decrease its velocity.	SAE level 0-1	-4.98	2.08	-9.00	-2.00
		SAE level 2-3	-5.21	2.46	-9.00	-2.50
		SAE level 4-5	-4.63	2.03	-9.00	-2.44
		SAE level 4-5 (aggressive mode)	-7.44	1.78	-9.00	-4.00
Normal deceleration (m/s²)	The maximum deceleration that a vehicle can use under normal driving conditions.	SAE level 0-1	-2.88	1.83	-7.50	-1.00
		SAE level 2-3	-4.00	2.26	-7.50	-2.00
		SAE level 4-5	-2.35	1.66	-7.50	-0.61
		SAE level 4-5 (aggressive mode)	-3.66	1.31	-4.50	-1.00
Reaction time (s)	The time it takes a vehicle to react to speed changes in the preceding vehicle.	SAE level 0-1	0.90	-	-	-
		SAE level 2-3	0.30	-	-	-
		SAE level 4-5	0.30	-	-	-
		SAE level 4-5 (aggressive mode)	0.20	-	-	-
Reaction time at traffic light (s)	The time takes a vehicle to react to restart at the green light.	SAE level 0-1	1.30	-	-	-
		SAE level 2-3	0.20	-	-	-
		SAE level 4-5	0.50	-	-	-
		SAE level 4-5 (aggressive mode)	0.10	-	-	-
Standstill distance (m)	The distance between stopped vehicles at the maximum traffic density.	SAE level 0-1	1.92	1.05	0.00	5.10
		SAE level 2-3	1.23	0.78	0.00	2.92
		SAE level 4-5	1.35	0.68	0.00	2.92
		SAE level 4-5 (aggressive mode)	1.25	0.53	0.50	2.00
Headway time (s)	The gap in seconds that a vehicle maintains.	SAE level 0-1	1.45	1.42	0.50	9.00
		SAE level 2-3	0.89	0.41	0.45	1.69
		SAE level 4-5	0.95	0.50	0.00	2.10
		SAE level 4-5 (aggressive mode)	0.66	0.20	0.42	1.00
Overtake speed threshold (%)	The percentage of the desired speed of a vehicle which decides to overtake.	SAE level 0-1	90.00	-	-	-
		SAE level 2-3	85.00	-	-	-
		SAE level 4-5	86.00	-	-	-
		SAE level 4-5 (aggressive mode)	85.00	-	-	-
Minimum lateral distance standing (m)	The minimum distance maintained by a passing vehicle when overtaking a stationary vehicle.	SAE level 0-1	0.60	0.40	0.20	1.00
		SAE level 2-3	0.15	0.00	0.15	0.15
		SAE level 4-5	0.45	0.30	0.15	0.75
		SAE level 4-5 (aggressive mode)	0.45	0.30	0.15	0.75
Aggressiveness level	A factor that reflects the aggressiveness of a vehicle that performs a lane change.	SAE level 0-1	-	-	0.60	0.63
		SAE level 2-3	-	-	0.21	0.45
		SAE level 4-5	-	-	0.10	1.00
		SAE level 4-5 (aggressive mode)	-	-	0.10	0.75
Cooperative lane change	When enabled, trailing vehicles execute change lanes to facilitate the lane change of a leading vehicle.	SAE level 0-1	no	-	-	-
		SAE level 2-3	yes	-	-	-
		SAE level 4-5	yes	-	-	-
		SAE level 4-5 (aggressive mode)	yes	-	-	-
Safety margin factor	The factor that determines when a vehicle can move at a priority junction.	SAE level 0-1	1.00	0.00	1.00	1.00
		SAE level 2-3	2.00	0.00	2.00	2.00
		SAE level 4-5	1.43	0.57	0.50	2.00
		SAE level 4-5 (aggressive mode)	1.47	0.64	0.88	2.50

Scenarios setting

- For integrating AVs to simulations, **fifteen scenarios** are selected in order multiple possible AV deployment scenarios to be investigated while including all four selected vehicle types.



Microscopic simulation

Traffic network geometric preparation

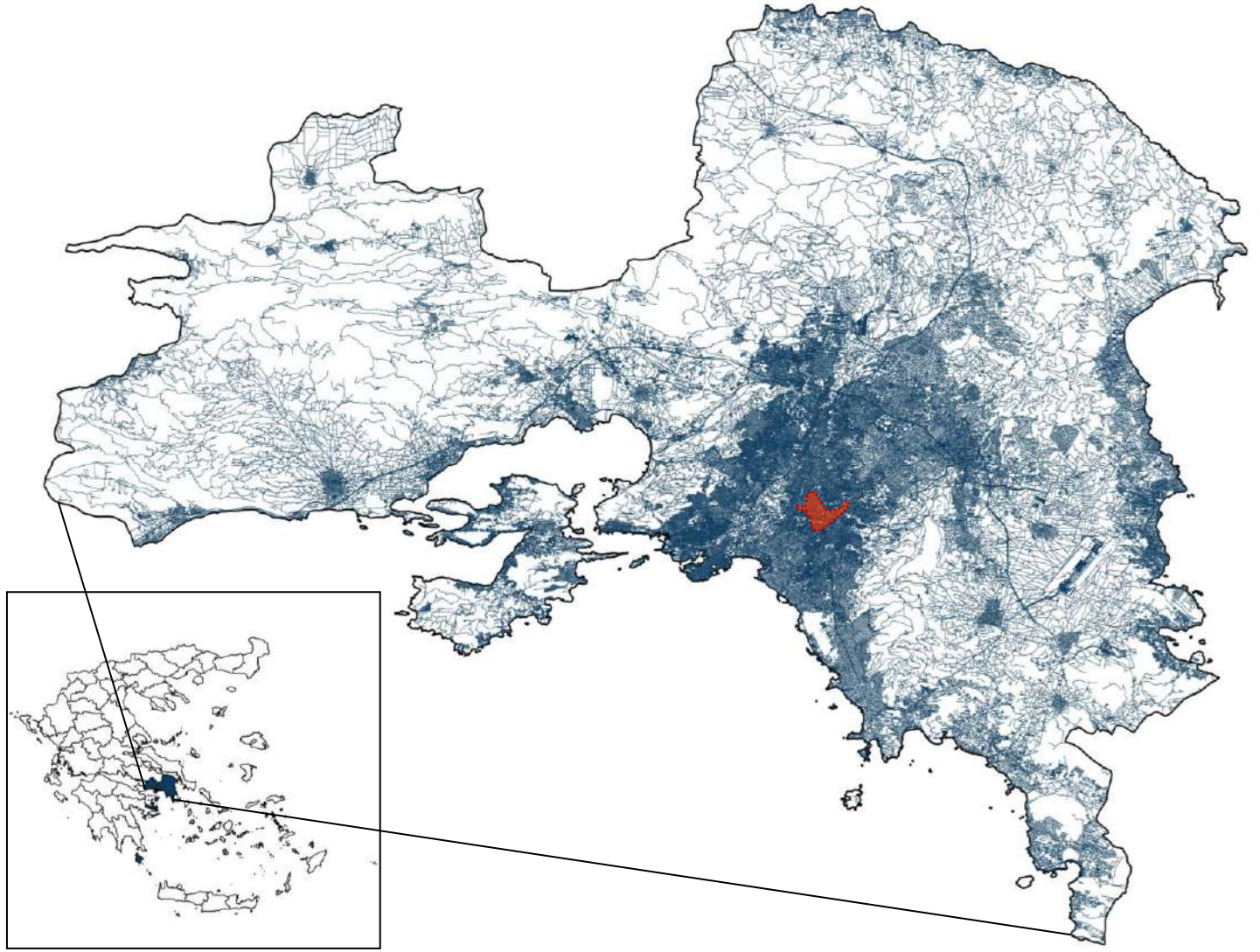
Traffic network calibration

Traffic network validation

Scenarios running

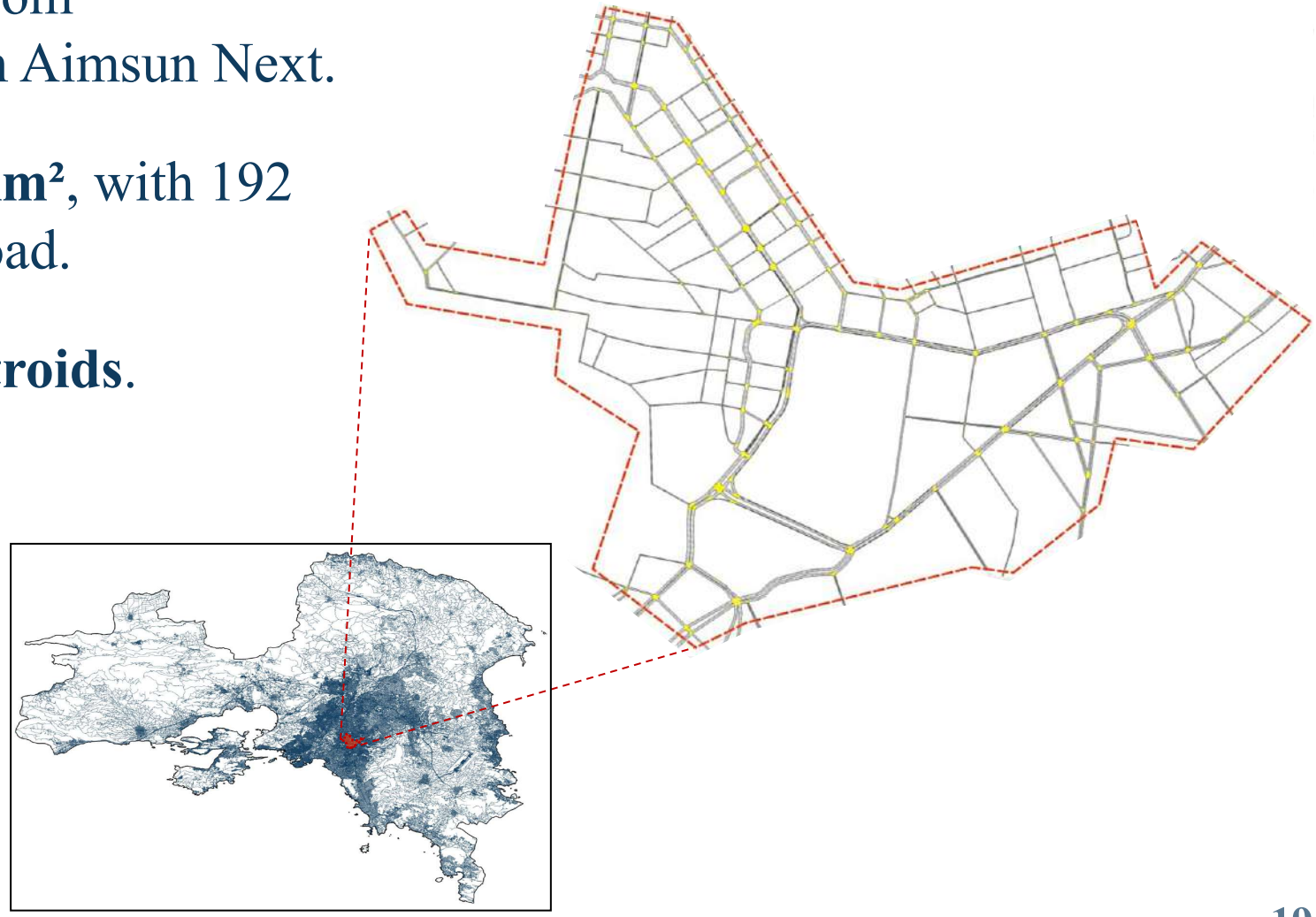
Study area

- The study area is the **Athens city centre**, characterized by dense traffic, high congestion, heavy motorbike and pedestrian flows and extensive public transport use.
- This area is considered as an **ideal testbed** for AV integration due to its complex, high-density and multimodal conditions.



Traffic network geometric preparation

- Network **geometry** was extracted from OpenStreetMap and implemented in Aimsun Next.
- The simulated network covers **3.2 km²**, with 192 nodes, 434 sections and 47 km of road.
- The OD matrix consists of **102 centroids**.
- **Public transport** is integrated, including 43 bus lines, 14 trolley lines, and 129 stops with frequencies and waiting times.

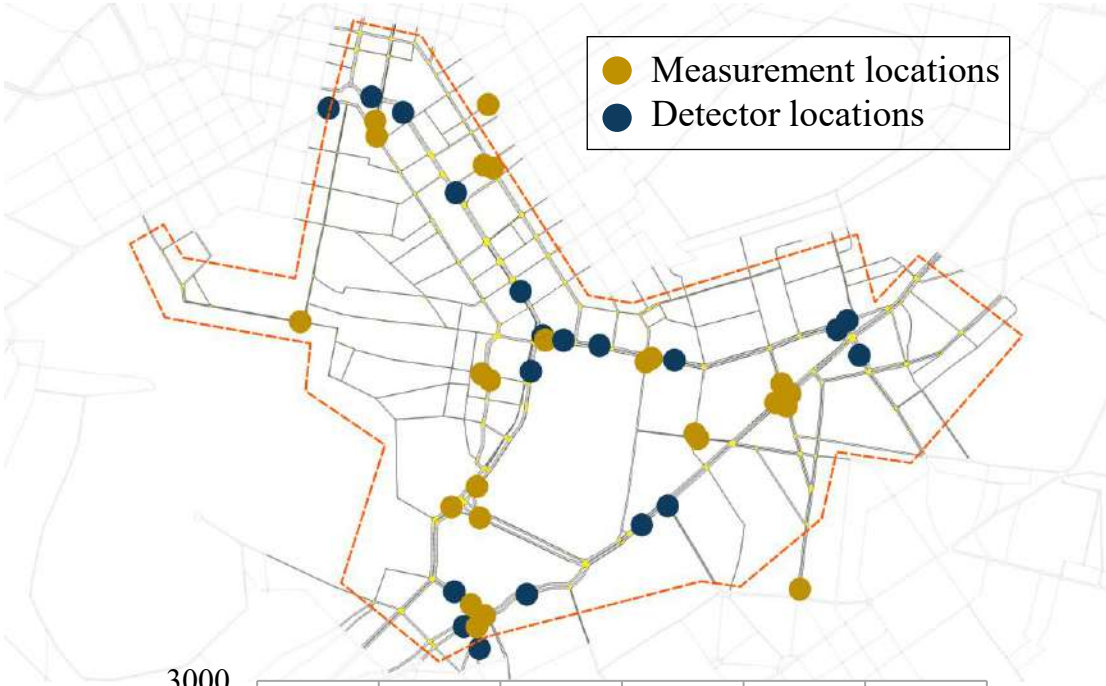


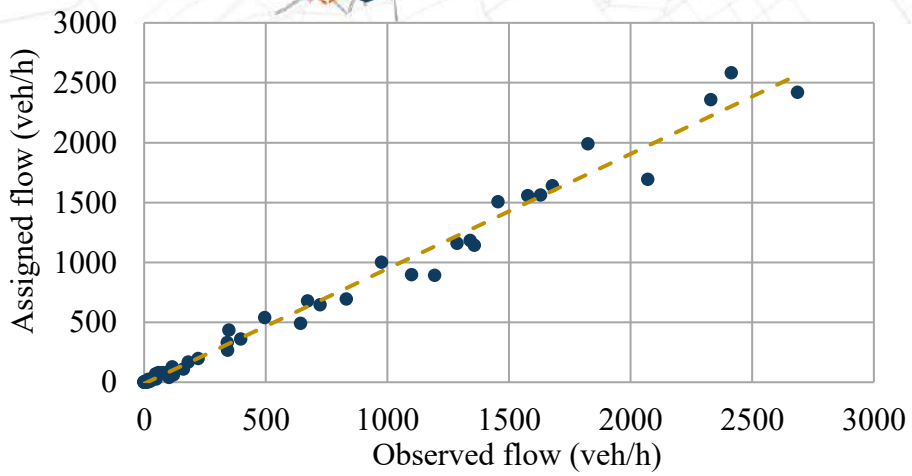
Traffic network calibration

- **Pedestrian flows** were integrated using manual counts from video recordings, capturing 6,912 pedestrians across visible crossings.
- Motorised **vehicle flows** for the morning peak (08:00-09:00) were obtained from 19 ATMC detectors.
- Additional **field measurements** captured vehicle fleet composition at key nodes, including motorbikes, heavy goods vehicles, passenger cars and taxis.

Calibration results & validation

- Calibration was successful, with **$R^2 = 0.98$** and GEH results meeting FHWA criteria (85.2% of links with GEH<5 and 100% with GEH<10).
- OD matrices were calibrated for 102 centroids, generating **20,987 peak-hour trips** (37% cars, 31% motorbikes, 28% taxis and 4% trucks).
- Model validation compared simulated and observed travel times (Google Maps API), showing **small deviations**.





y = 0.959x - 11.888
R^2 = 0.9837

Surrogate Safety Assessment

Derivation of surrogate safety measures

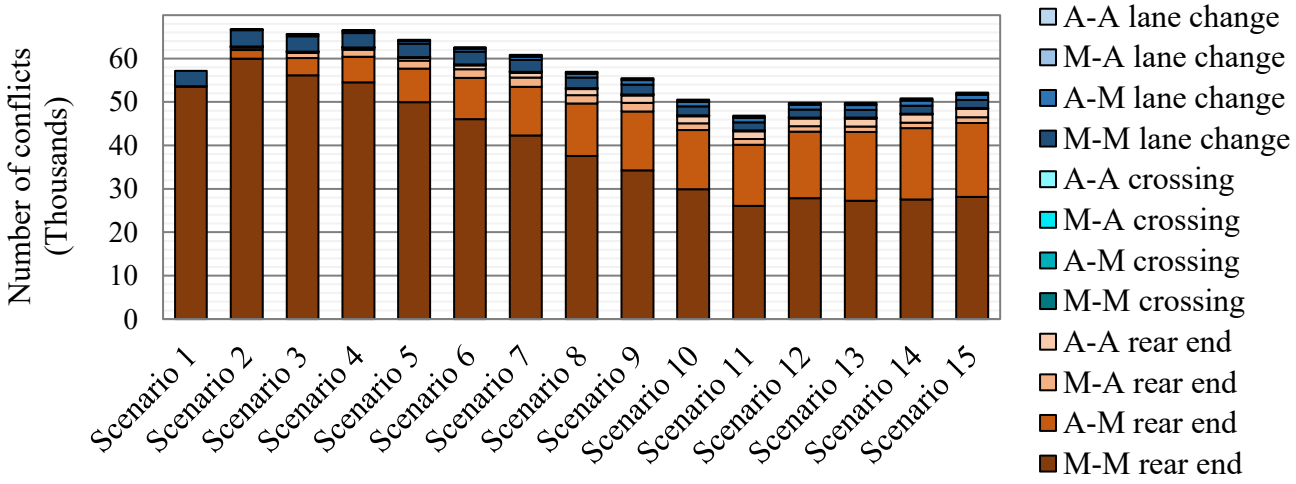
Convert traffic conflicts to crash risk

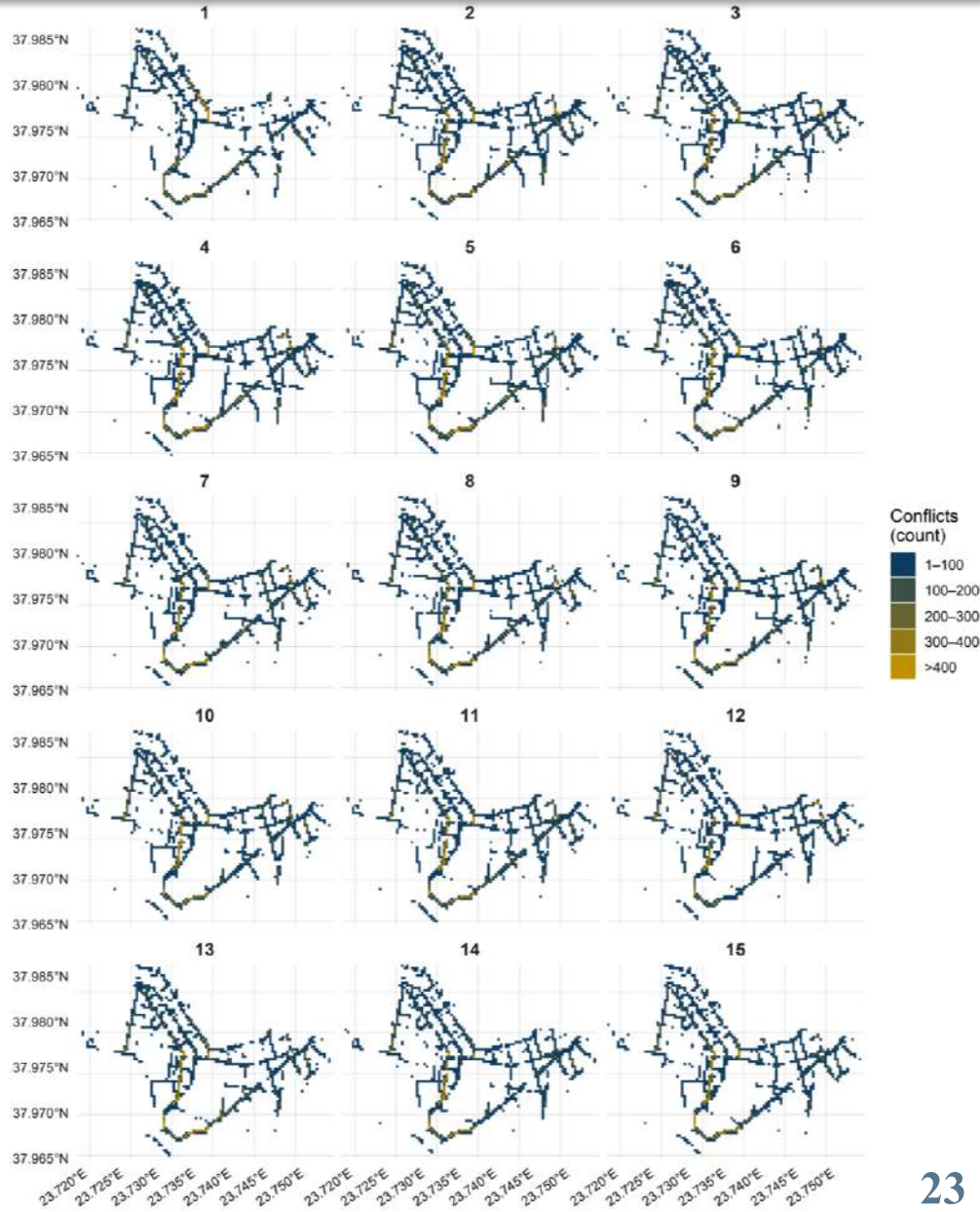
Crash risk validation

Employing field crash data

Derivation of surrogate safety measures

- **Vehicle trajectories** were extracted from the microsimulation at 0.1-second intervals and analysed using SSAM.
- Conflicts were identified using TTC thresholds, the default value (1.5 s) was **adapted for AVs** by SAE level (SAE 2-3: 0.9 s; SAE 4-5: 1.0 s; SAE 4-5 aggressive: 0.7 s).



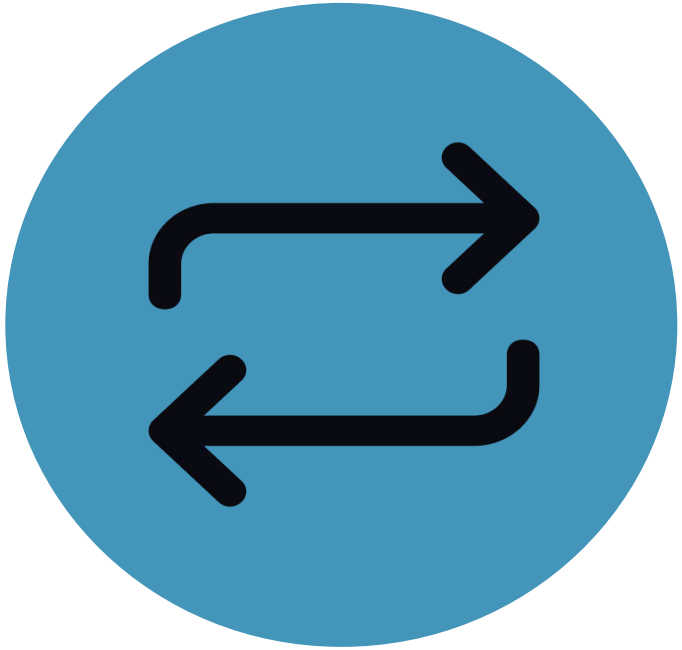


Convert traffic conflicts to crash risk

RQ2

- Earlier studies relied on simple crash-conflict ratios, while later work[1] introduced a **TTC-based formulation** to translate aggregated simulated conflicts into expected crash counts.

- This formulation is applied at the **individual conflict-event level**, enabling estimation of crash risk for each conflict and extended with automation-specific TTC thresholds to capture behavioural differences across vehicle types:



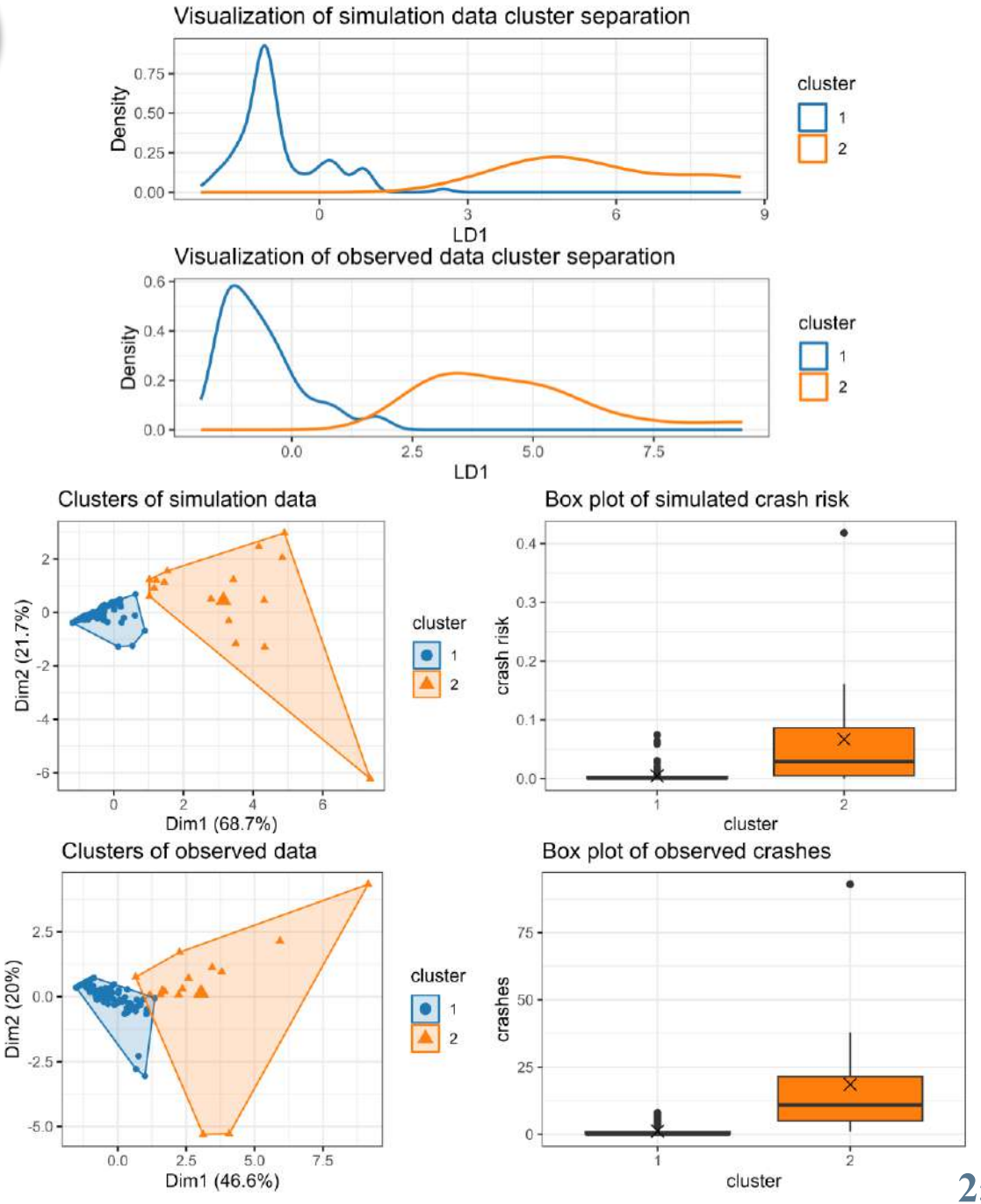
$$\text{Crash risk} = 2^{\frac{\log(0.5)}{\log\left(1+\frac{TTC_t - TTC_i}{TTC_i}\right)}}$$

- TTC_i: TTC value for the i-th observed conflict
- TTC_t: TTC threshold of the following vehicle in that conflict

[1] *Tarko, A.P., 2018. Estimating the expected number of crashes with traffic conflicts and the Lomax Distribution – A theoretical and numerical exploration. Accid. Anal. Prev. 113 January , 63–73. doi:10.1016/j.aap.2018.01.008*

Crash risk validation

- Crash risk estimates were further validated against observed crash data using **K-means clustering**.
- Observed **crash data** for the study area were sourced from ELSTAT and include road-level injury crashes from 2017-2019.
- The Elbow method indicated two clusters in both datasets, representing **low-risk and high-risk roads**.
- LDA was used to assess how well the identified clusters are separated, indicating clear separation with **no overlap**.



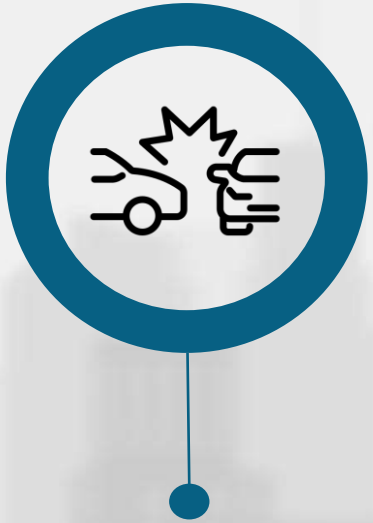
Spatial validation results

- Simulated and observed risk classifications agreed for **87.7% of roads** (107 of 122).
- **Geospatial maps** are used to identify spatial agreement and mismatches.
- Misclassifications (15 roads) are associated with locations **lacking traffic flow data used for calibration**, emphasizing the importance of data availability and calibration quality for reliable safety estimation.

Crash risk assessment

Road-level analysis

Focuses on identifying key traffic, infrastructure and automation-related factors associated with crash risk and analysing their influence.

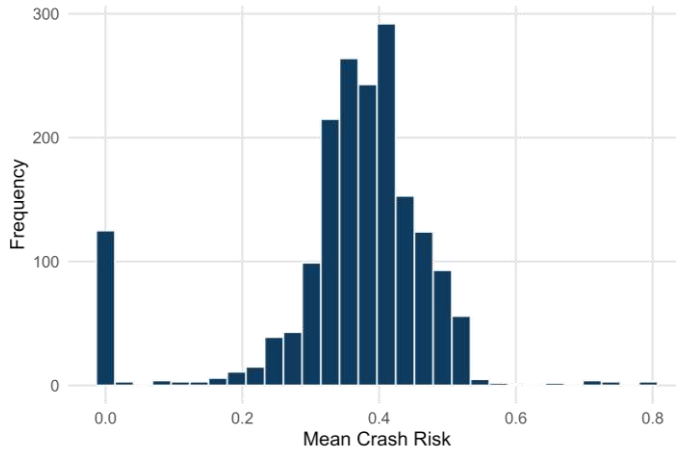


Conflict-level analysis

Investigates the spatial dynamics of crash risk, identifying and modelling crash hotspots.

Road-level safety evaluation: Crash risk factor analysis

- Road-level crash risk was analysed comprising over **8.8 M conflict events** aggregated to 1,812 roads.
- Exploratory analysis revealed non-normality, non-linear relationships and multicollinearity.
- Linear Regression, Random Forest, and XGBoost were evaluated using standard error metrics (RMSE, MAE, R^2).
- **XGBoost achieved the best performance** (R^2 = 0.73, RMSE = 0.18, MAE = 0.10) and was selected for crash risk factor analysis.
- **SHAP values** were used to rank the relative influence of variables in the final XGBoost model.

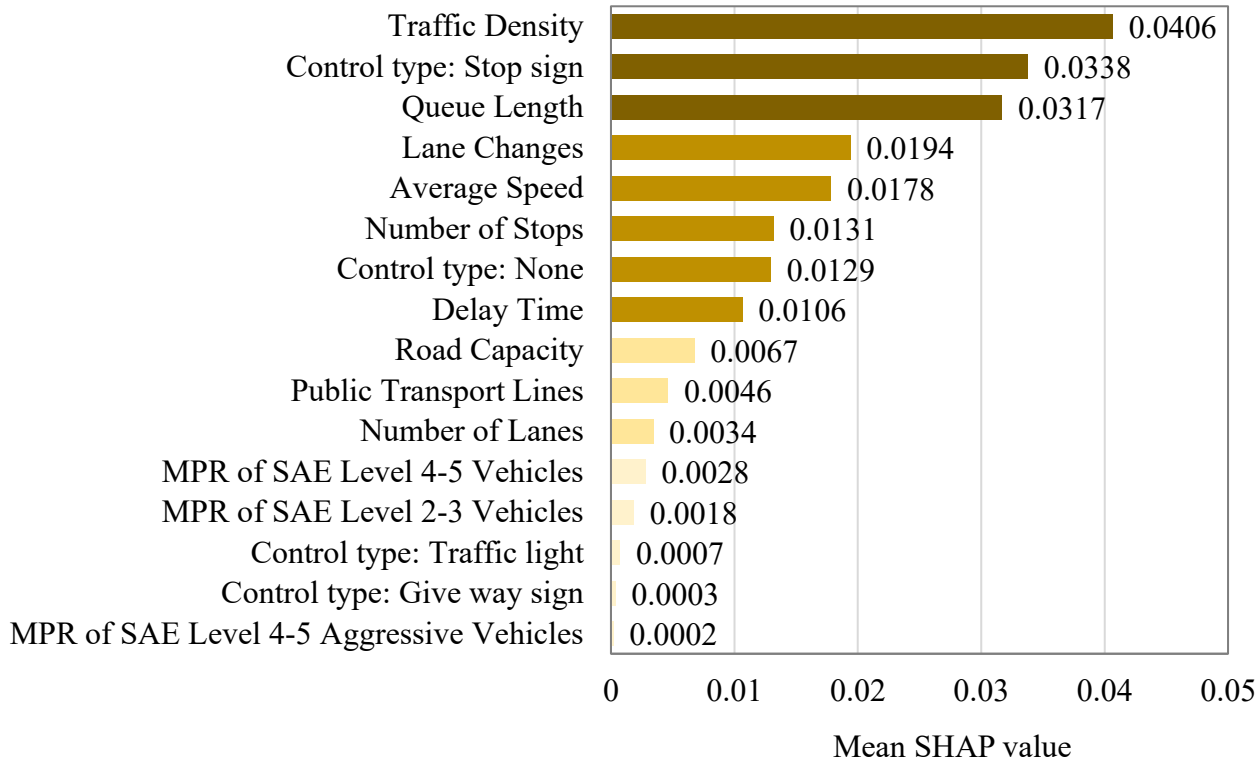


Model	RMSE	MAE	R^2
XGBoost (tuned)	**0.1757**	**0.0950**	**0.7331**
Random Forest (tuned)	0.2013	0.0489	0.6498

Crash risk factors ranking

RQ3

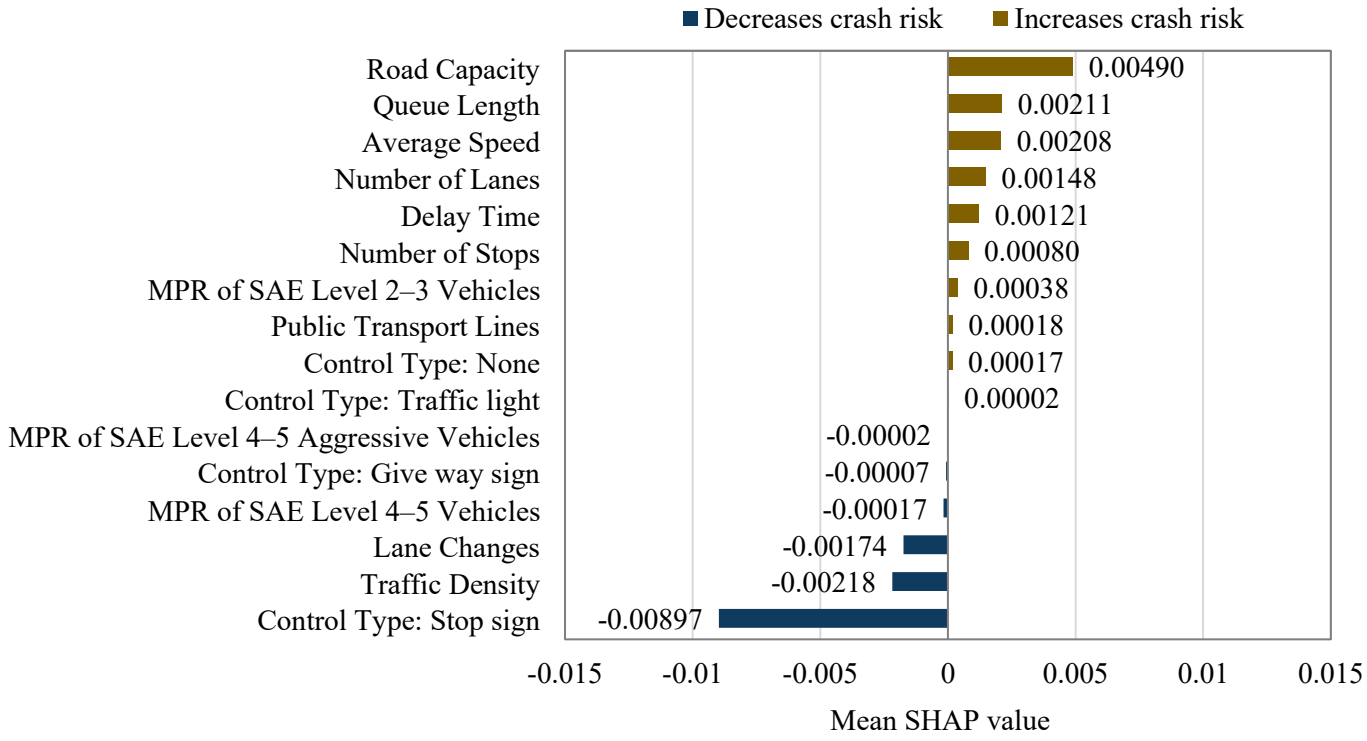
- Traffic-related factors remain the **main contributors to crash risk** even with automated vehicles present.
- Operational instability indicators exhibit **moderate influence**, highlighting the role of traffic dynamics in safety outcomes.
- Automation-related variables contribute indirectly, with SAE Level 4-5 penetration **more influential** than SAE 2-3, while the aggressive SAE 4-5 variant shows minimal impact.

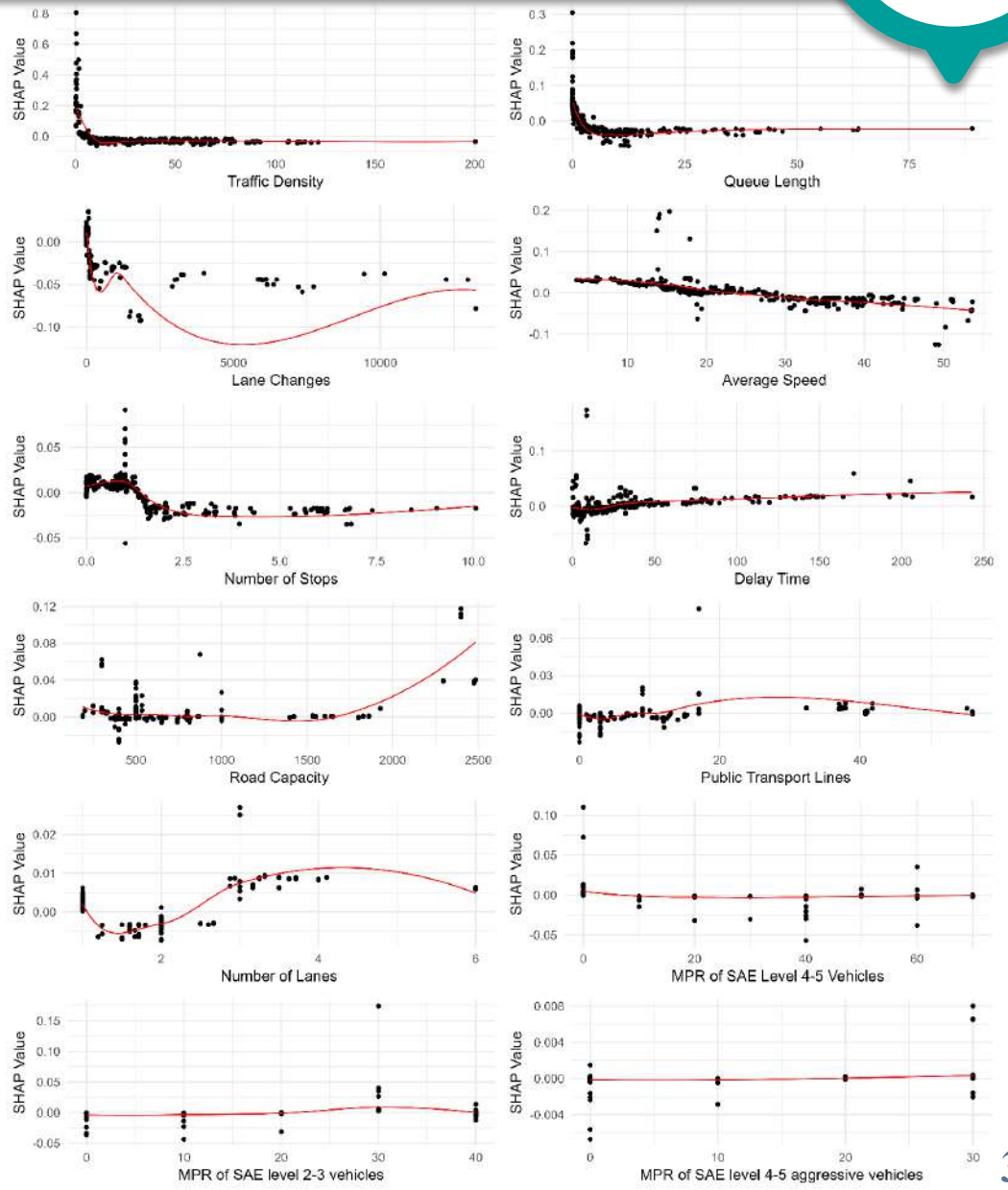


Impacts of factors on crash risk

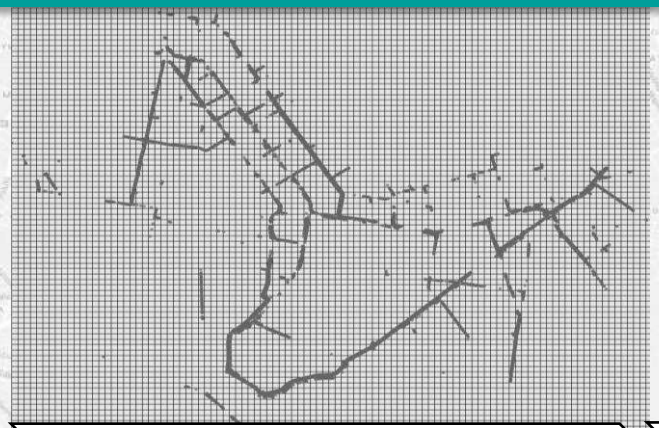
RQ4

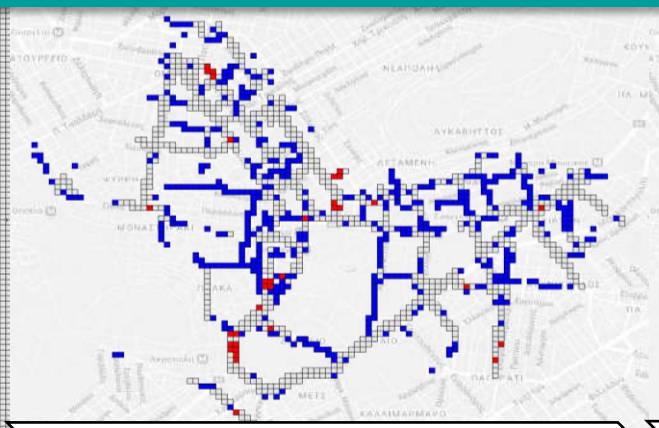
- **Automation effects are asymmetric**: SAE Level 4-5 penetration reduces crash risk, while higher SAE Level 2-3 penetration increases risk, aggressive SAE Level 4-5 shows minimal influence.

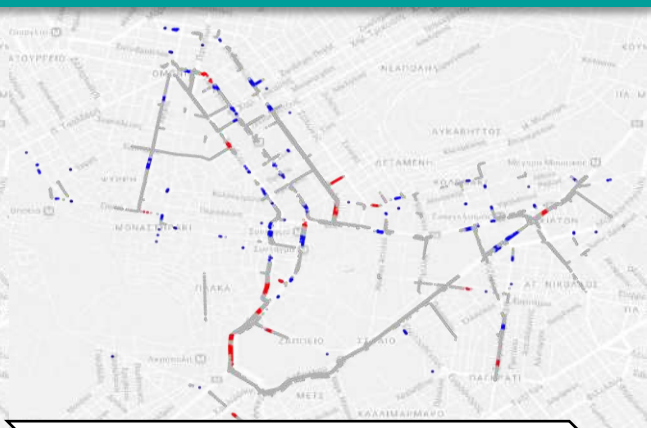




Conflict-level safety evaluation: Crash hotspot dynamics







Conflicts events identification

- A representative simulation replication was selected to ensure comparability across scenarios, resulting in over 880,000 conflict events.

Grid-based conflicts aggregation

- Overlay a fixed spatial grid (30m×30m).
- Assign conflict events to its corresponding grid cell.
- Aggregate crash risk at the grid-cell level.

Hotspot classification of grid cells

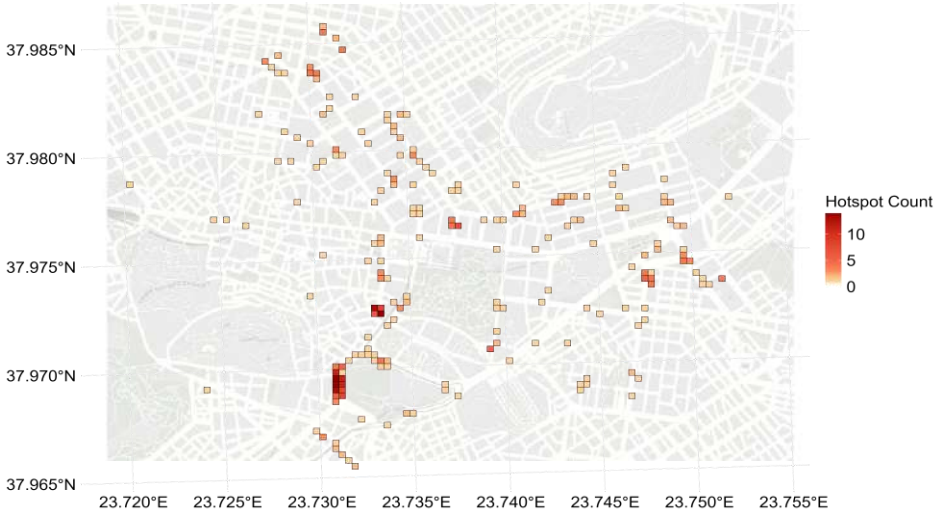
- Classify grid cells into Hotspot (high-risk, Gi*>1.96), Coldspot (low-risk, Gi*<-1.96), or Neutral (no significant clustering).

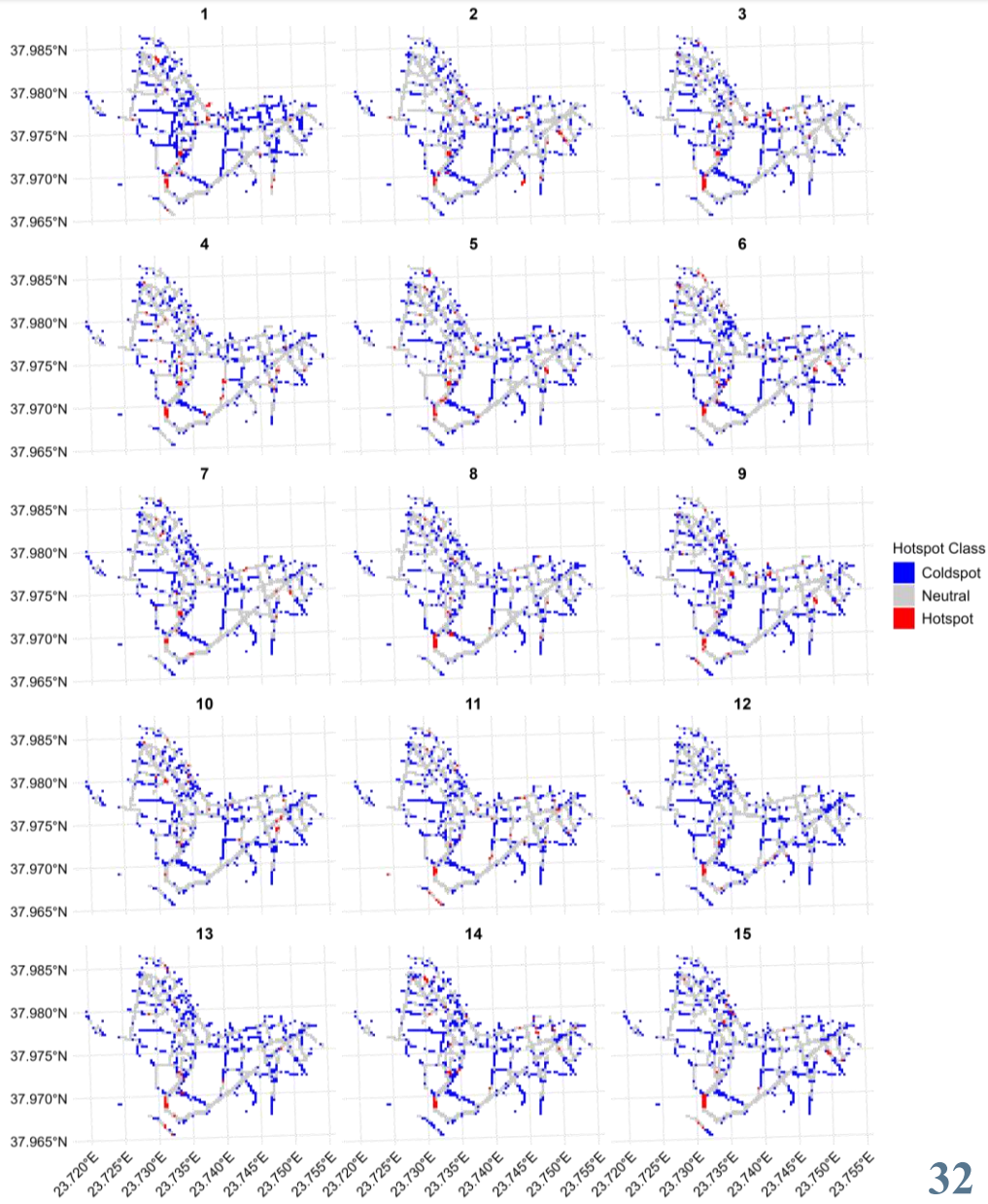
Conflict events by hotspot status

- Associate each conflict with the hotspot class of its corresponding grid cell.
- Fit a model to predict the probability that a conflict occurs within a hotspot area.

Crash hotspot detection approach

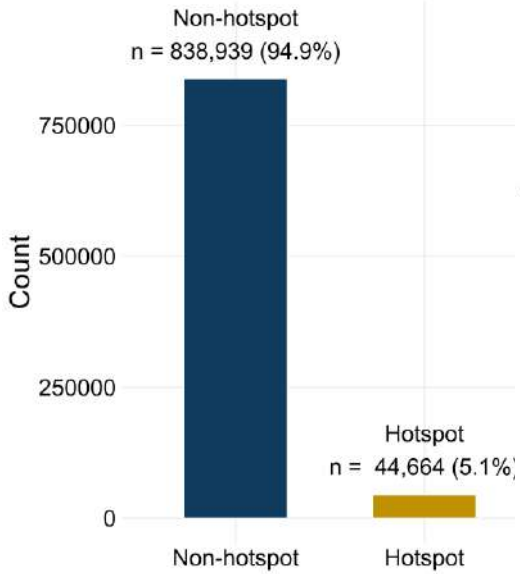
- The **Getis-Ord Gi* statistic** was applied to quantify local spatial autocorrelation, comparing each location and its neighbours to the network mean to identify statistically significant hotspots and coldspots.
- Hotspot extent varies across AV scenarios, with overall **reductions** relative to the baseline, while persistent hotspots remain at **complex network locations**.

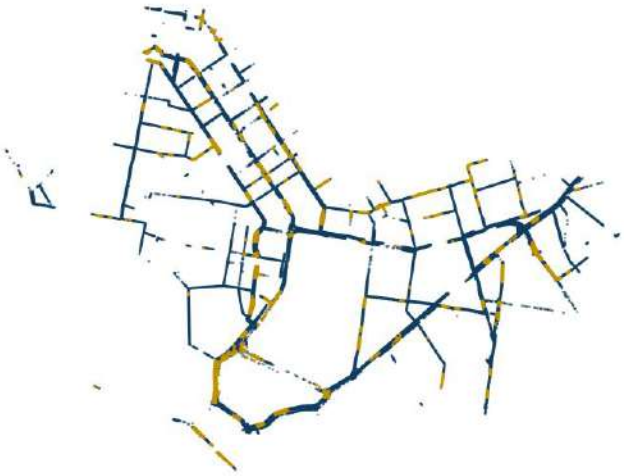


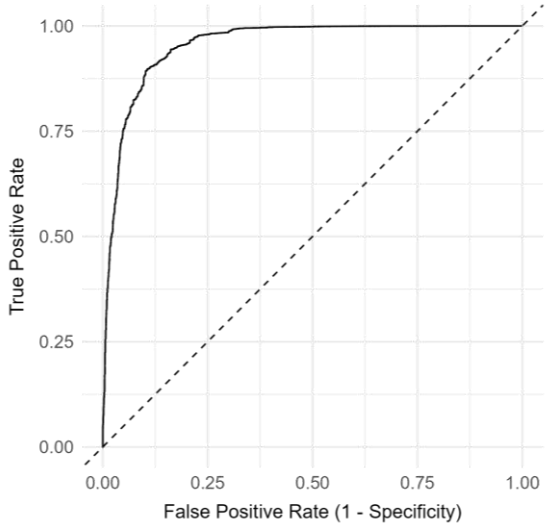


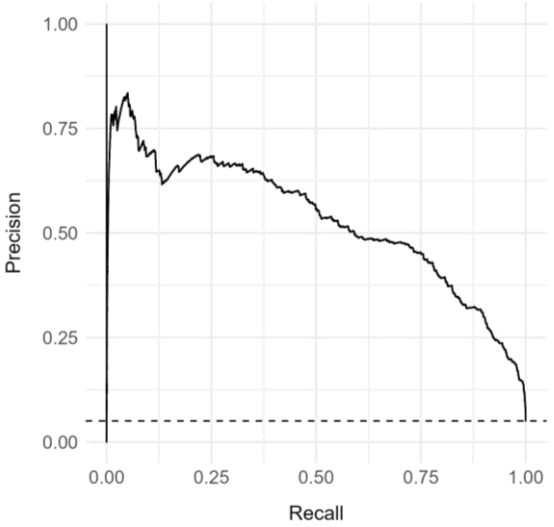
Crash hotspot probability modelling

- Binary **hotspot occurrence** (hotspot = 1, non-hotspot = 0) exhibited strong class imbalance, requiring **oversampling** (≈1.6 M conflict records).
- Logistic, spatial logistic, GAM and tree-based ML models (Random Forest, XGBoost) were **tested and compared** for crash risk modelling.
- The **binomial GAM** was selected, capturing nonlinear and spatial effects (AUC = 0.96, Adjusted $R^2 \approx 0.68$).









Parametric effects on hotspot formation

- **Intersections and crossing conflicts** exhibit a significantly higher likelihood of hotspot formation.
- Traffic control reduces hotspot probability, with both **stop signs and traffic lights** associated with lower risk.
- Conflicts involving **AV-human and AV-AV pairings** show higher hotspot probability.
- Crash hotspot probability initially increases at low-to-moderate AV penetration but declines as **automation becomes dominant**.

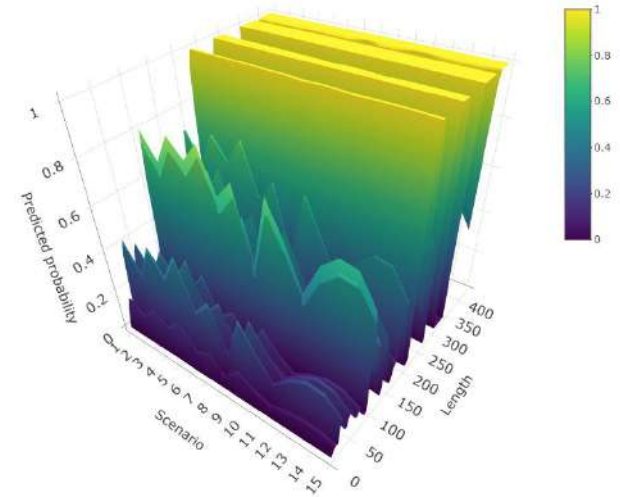
Parametric coefficients of the GAM (Response variable: hotspot occurrence)

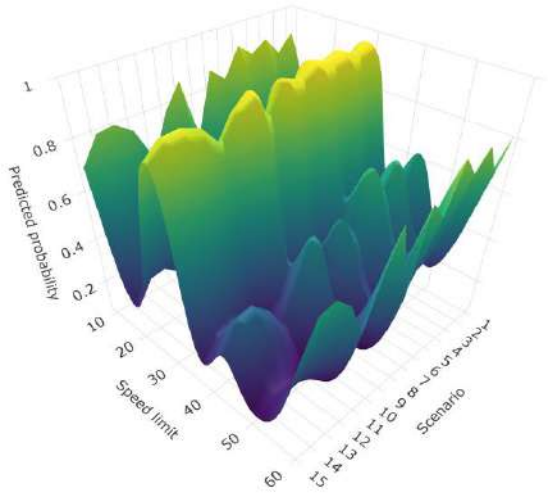
Predictor	Estimate	Std. Error	z value	p-value	Sig.
Intercept	-15.615	0.255	-61.160	<0.001	***
Road context: intersection *(Ref.: link)*	0.490	0.075	6.575	<0.001	***
Conflict type: lane change *(Ref.: crossing)*	-0.674	0.062	-10.788	<0.001	***
Conflict type: rear end *(Ref.: crossing)*	-0.647	0.061	-10.572	<0.001	***
Control: stop *(Ref.: none)*	-0.705	0.055	-12.711	<0.001	***
Control: traffic lights *(Ref.: none)*	-1.147	0.026	-43.966	<0.001	***
Vehicle pair: AV-manual *(Ref.: manual-manual)*	0.022	0.008	2.644	0.008	**
Vehicle pair: manual-AV *(Ref.: manual-manual)*	0.107	0.019	5.667	<0.001	***
Vehicle pair: AV-AV *(Ref.: manual-manual)*	0.098	0.023	4.267	<0.001	***
Scenario: 2 *(Ref.: 1)*	-0.592	0.017	-34.432	<0.001	***
Scenario: 3 *(Ref.: 1)*	0.114	0.017	6.709	<0.001	***
Scenario: 4 *(Ref.: 1)*	-0.321	0.017	-18.350	<0.001	***
Scenario: 5 *(Ref.: 1)*	0.222	0.017	13.184	<0.001	***
Scenario: 6 *(Ref.: 1)*	-0.777	0.018	-42.534	<0.001	***
Scenario: 7 *(Ref.: 1)*	-0.372	0.018	-20.679	<0.001	***
Scenario: 8 *(Ref.: 1)*	-1.745	0.019	-89.766	<0.001	***
Scenario: 9 *(Ref.: 1)*	-0.372	0.018	-20.393	<0.001	***
Scenario: 10 *(Ref.: 1)*	-1.043	0.020	-53.369	<0.001	***
Scenario: 11 *(Ref.: 1)*	-1.818	0.021	-85.199	<0.001	***
Scenario: 12 *(Ref.: 1)*	-1.184	0.021	-55.630	<0.001	***
Scenario: 13 *(Ref.: 1)*	-0.909	0.020	-45.913	<0.001	***
Scenario: 14 *(Ref.: 1)*	-0.961	0.020	-48.134	<0.001	***
Scenario: 15 *(Ref.: 1)*	-1.488	0.020	-76.235	<0.001	***

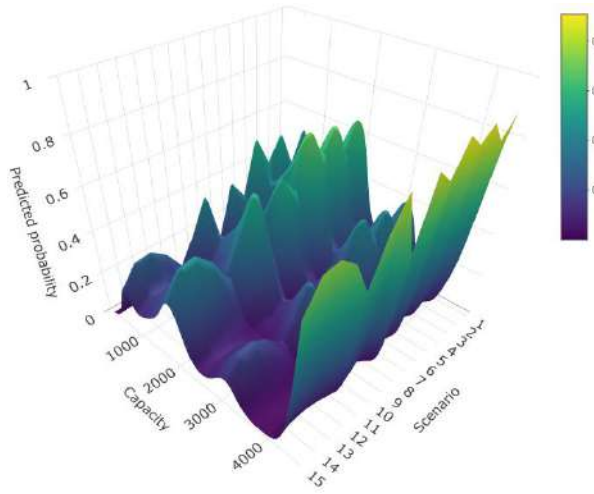
Nonlinear effects roadway characteristics

RQ4

- Automation leads to the **largest reductions** in hotspot probability on short segments with higher speed limits and limited road capacity.
- Long, low-speed, or high-capacity segments show relatively **stable hotspot probabilities** across scenarios, indicating limited sensitivity to automation effects.







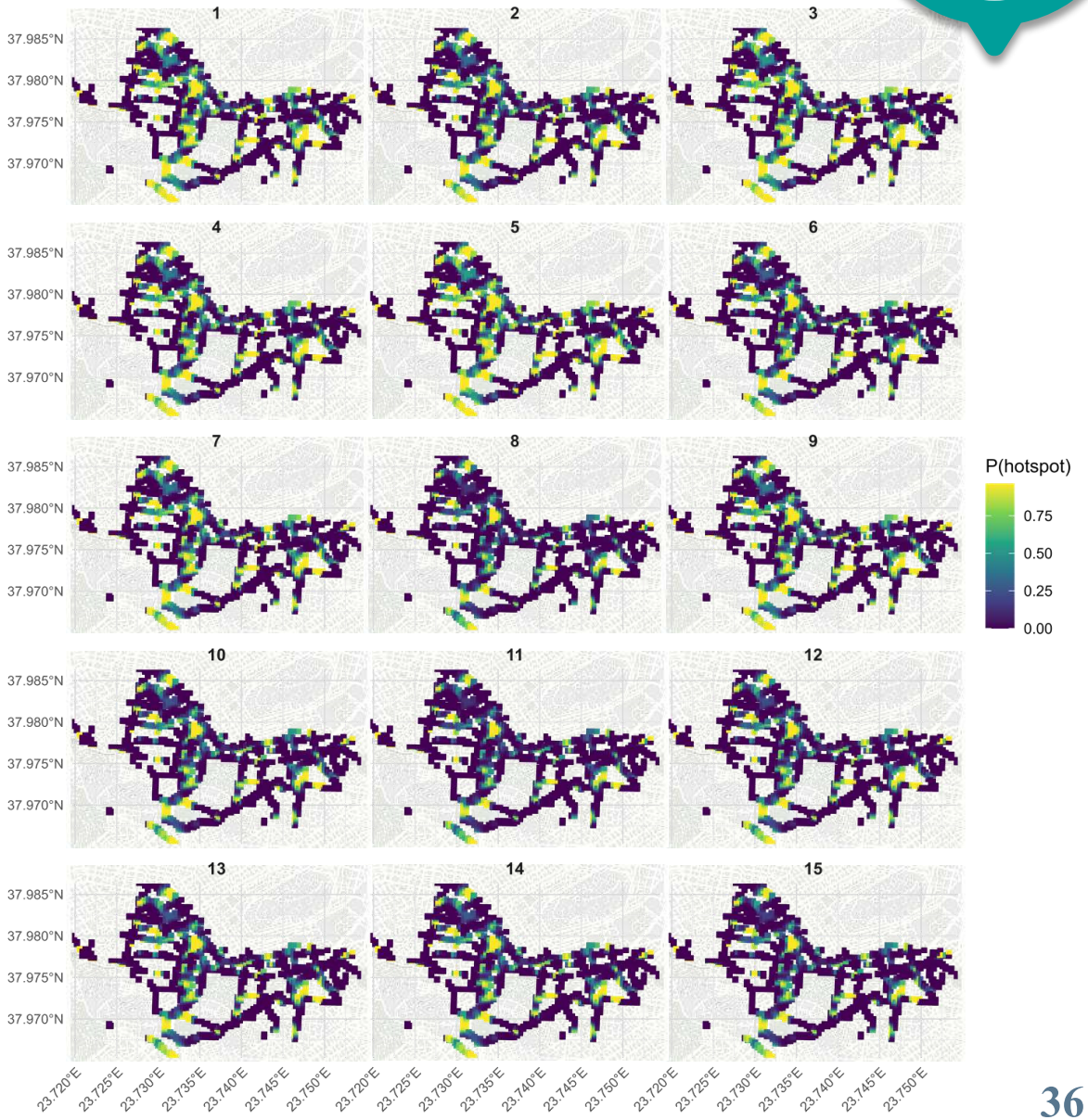
Smooth terms of the GAM (Response variable: hotspot occurrence)

Smooth term	EDF	Ref.df	Chi.sq	p-value	Sig.
s(x, y)	196.25	199	131,089	<0.000	***
s(Length)	28.96	29	12,714	<0.000	***
s(Speed limit)	6.95	7	2,679	<0.001	***
s(Capacity)	8.97	9	3,377	<0.001	***

Spatial visualization

RQ4

- Early scenarios show widespread hotspots along **arterials and intersections**, while higher automation levels lead to fewer hotspots.
- Persistent hotspots remain at **complex intersections** and high-demand corridors, suggesting that infrastructure and control constraints continue to dominate risk, even under high automation.



Summary of key findings (1/2)

- A validated **simulation-based framework** links behavioural modelling, surrogate safety measures and observed crash patterns, overcoming the lack of AV crash data.
- Automation affects crash risk **gradually and context-dependently**, shaped by traffic mix, infrastructure and behavioural variability.
- Traffic and infrastructure factors remain the **strongest contributors** to crash risk, even under increasing automation.
- Crash risk and hotspot formation change with increasing AV penetration, showing **nonlinear** responses during the transition.

Summary of key findings (2/2)

RQ5

Crash risk evolution across automation phases:

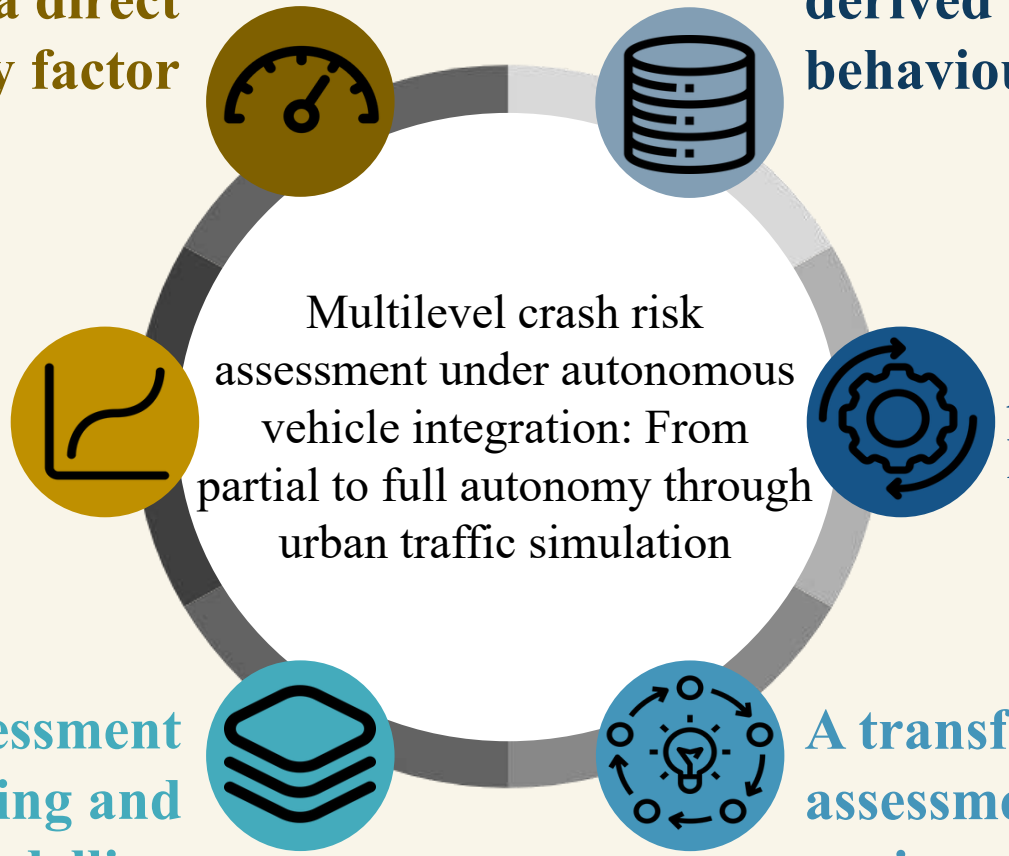
Early and moderate automation	Mixed transition	High automation
Crash risk increases with rising SAE 2-3 penetration.	Scenario effects become negative, indicating reduced crash risk.	Substantial crash risk reduction is observed at high automation levels.
Higher hotspot probability is observed for mixed manual-AV and AV-manual interactions.	SAE 4-5 vehicles introduce stabilising effects, while SAE 2-3 vehicles still introduce variability.	Consistent risk-reducing effects of SAE 4-5 penetration are evident, with minimal influence from aggressive AVs.
Hotspots spread spatially.	Hotspots consolidate but persist in complex locations.	Hotspots become few and less intense, remaining mainly in complex areas.
Scenarios 2-5, 10-40% AV MPR, SAE 0-3	**Scenarios 6-10, 50-90% AV MPR, mix of SAE 2-3, SAE 4-5 and aggressive AVs**	**Scenarios 11-15, 100% AV MPR, mix of SAE 2-3, SAE 4-5 and aggressive AVs**

Innovative contributions

Automation acts as a system stabiliser rather than a direct safety factor

A comprehensive, systematically derived repository of AV behavioural parameters

Nonlinear evolution of crash risk across automation levels



A novel TTC-based event-level conversion framework

A multiscale safety assessment combining machine learning and spatial-statistical modelling

A transferable methodology for safety assessment in emerging mobility environments

Practical and policy implications

- Researchers can use the **AV parameter repository** to improve consistency in microsimulation studies.
- Planners and analysts can anticipate **how crash risk evolves** across AV market penetration levels using the proposed framework.
- Regulators gain a **validated tool for safety evaluation** in data-poor, high-automation scenarios where crash records do not yet exist.
- Infrastructure and traffic management decisions can be informed by identifying conditions **where mixed traffic increases risk** and where automation delivers the largest safety gains.

Limitations

- AV behaviour is simulated using models, which were originally developed **for human-driven vehicles**.
- AV behavioural parameters are mostly derived from prototype trials, small-scale experiments and prior simulations, with **limited evidence** from large-scale field deployment, particularly for SAE Level 4-5.
- **Grouping SAE levels** (0-1, 2-3, 4-5) was necessary due to limited data and may hide behavioural differences between levels.
- The analysis focuses exclusively on **vehicle-vehicle interactions**, excluding vulnerable road users.

Future research directions

- Examine **manufacturer-specific** behaviours, supported by sensitivity analyses of key behavioural parameters.
- Validate the framework using emerging **AV crash datasets**, especially for high SAE 4-5 penetration.
- Broaden applicability by incorporating additional **vehicle classes, VRU types** and diverse network types.
- Enrich calibration and validation of simulation model by incorporating **behavioural and contextual data** (e.g., rule violations, telematics-based indicators).
- Advance spatiotemporal risk modelling to capture how crash risk evolves **jointly over time and space**.

Department of Transportation Planning and Engineering
School of Civil Engineering
National Technical University of Athens

Thesis submitted for the degree of
Doctor of Philosophy

22 January 2026

Multilevel crash risk assessment under autonomous vehicle integration: From partial to full autonomy through urban traffic simulation

Maria G. Oikonomou

Advisory Committee:

Prof. George Yannis (Supervisor)
Prof. Eleni Vlahogianni
Asst. Prof. Konstantinos Gkiotsalitis