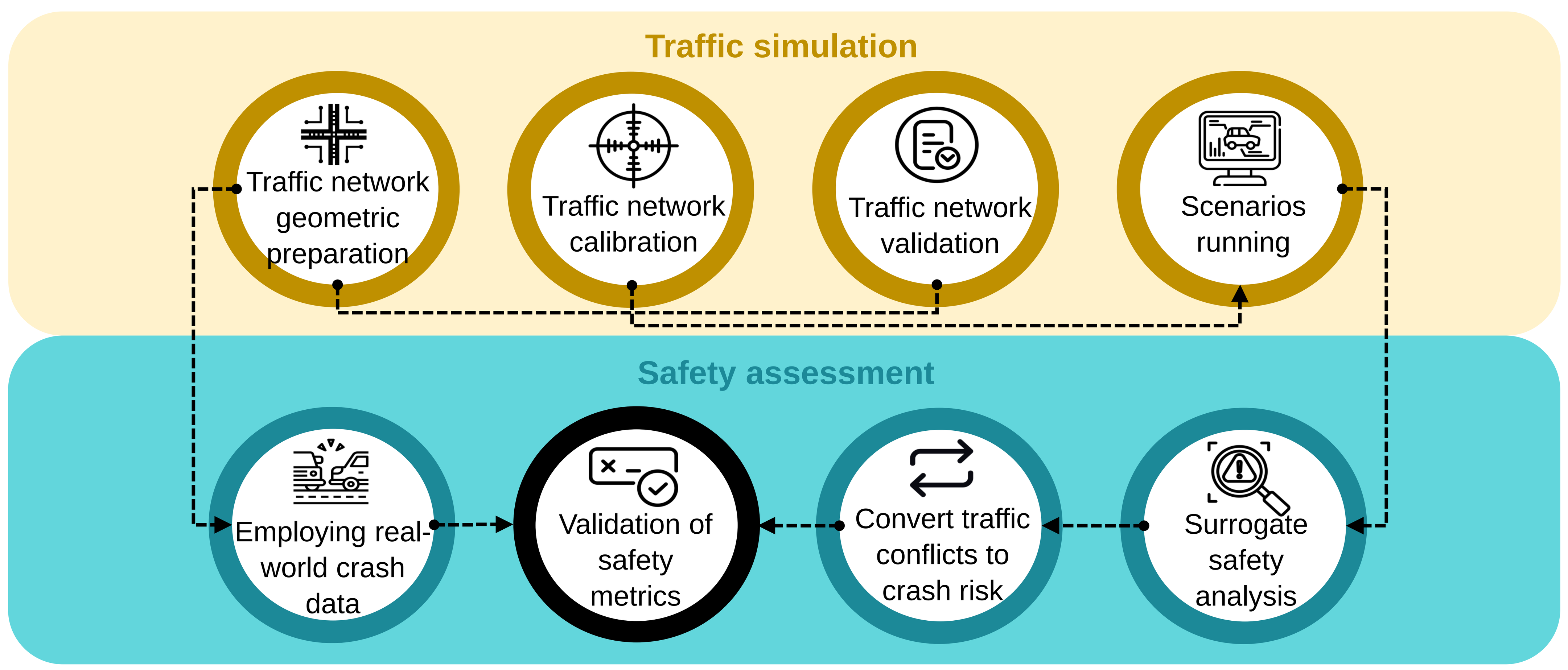
Validation Analysis of Traffic Simulation Safety Metrics with Real-World Crash Data

Maria G. Oikonomou and George Yannis

Department of Transportation Planning and Engineering
National Technical University of Athens

TRBAM-25-02287

METHODOLOGICAL FRAMEWORK



INTRODUCTION

Traffic simulation is extensively employed in transportation engineering, particularly for safety analysis. The application of microscopic simulation has advanced due to the development of methods and tools to quantify surrogate safety measures. The conflict-based approach is a widespread method for identifying high-risk locations without requiring field crash data. However, a critical limitation remains the **absence of comprehensive models for simulating crashes and the investigation of safety outcomes validity**.

Objective

This study aims to **bridge the gap between simulation models and real-world safety observations**, contributing to the advancement of more robust safety assessment methodologies. It presents a comprehensive comparative analysis of traffic safety metrics derived from both simulated and real-world data, employing clustering technique to identify safety patterns.

METHODOLOGY

Study area

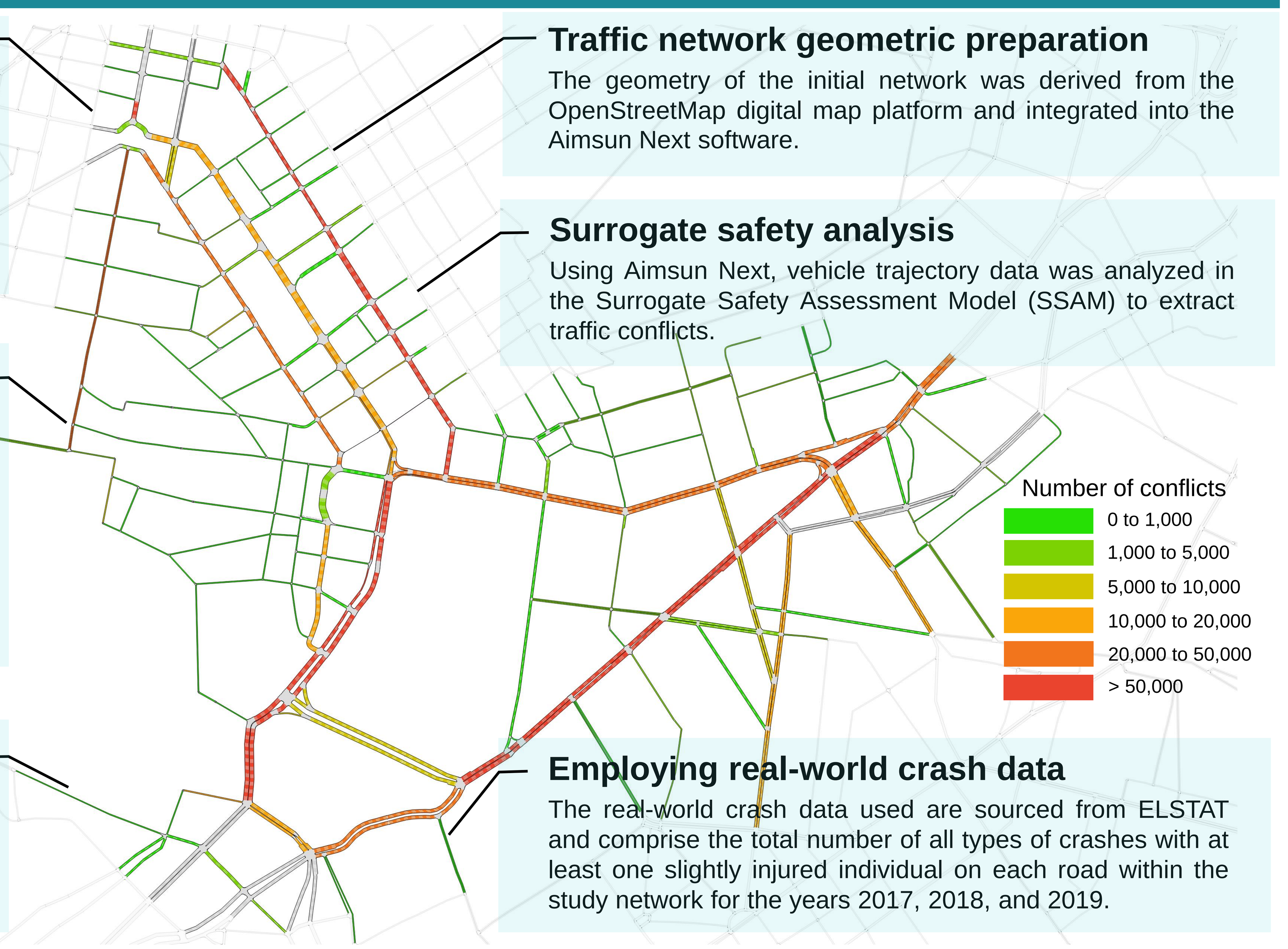
The study area is the **city center of Athens, Greece**, a critical urban zone characterized by dense motorized traffic covering approximately 3.2 square kilometers.

Traffic network calibration

- Detector recording flows integration
- Field measurement flows integration
- Camera-derived pedestrian flows integration

Traffic network validation

Derivation of travel time and speed using Google Maps API



Traffic network geometric preparation

The geometry of the initial network was derived from the OpenStreetMap digital map platform and integrated into the Aimsun Next software.

Surrogate safety analysis

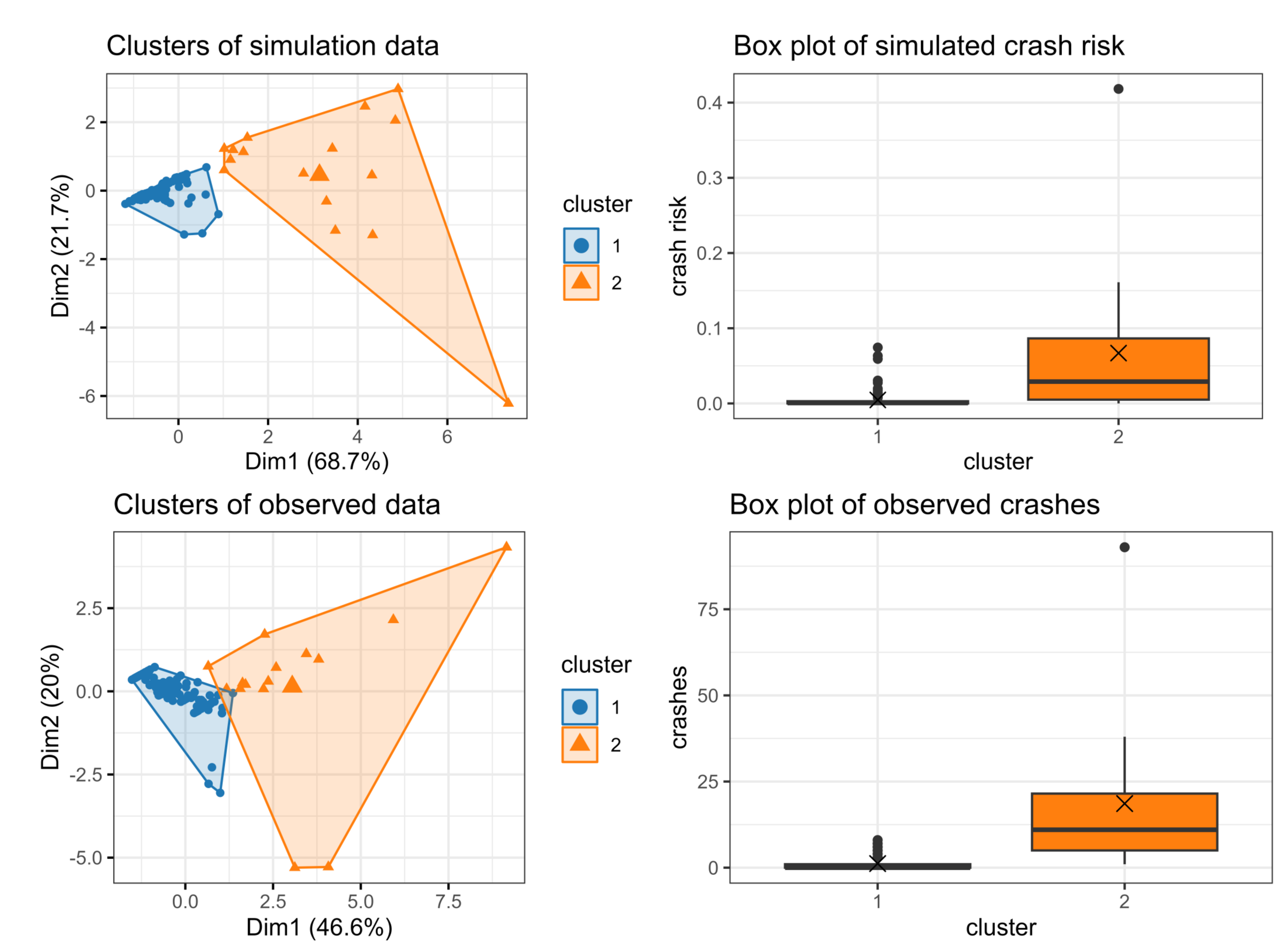
Using Aimsun Next, vehicle trajectory data was analyzed in the Surrogate Safety Assessment Model (SSAM) to extract traffic conflicts.

Employing real-world crash data

The real-world crash data used are sourced from ELSTAT and comprise the total number of all types of crashes with at least one slightly injured individual on each road within the study network for the years 2017, 2018, and 2019.

RESULTS

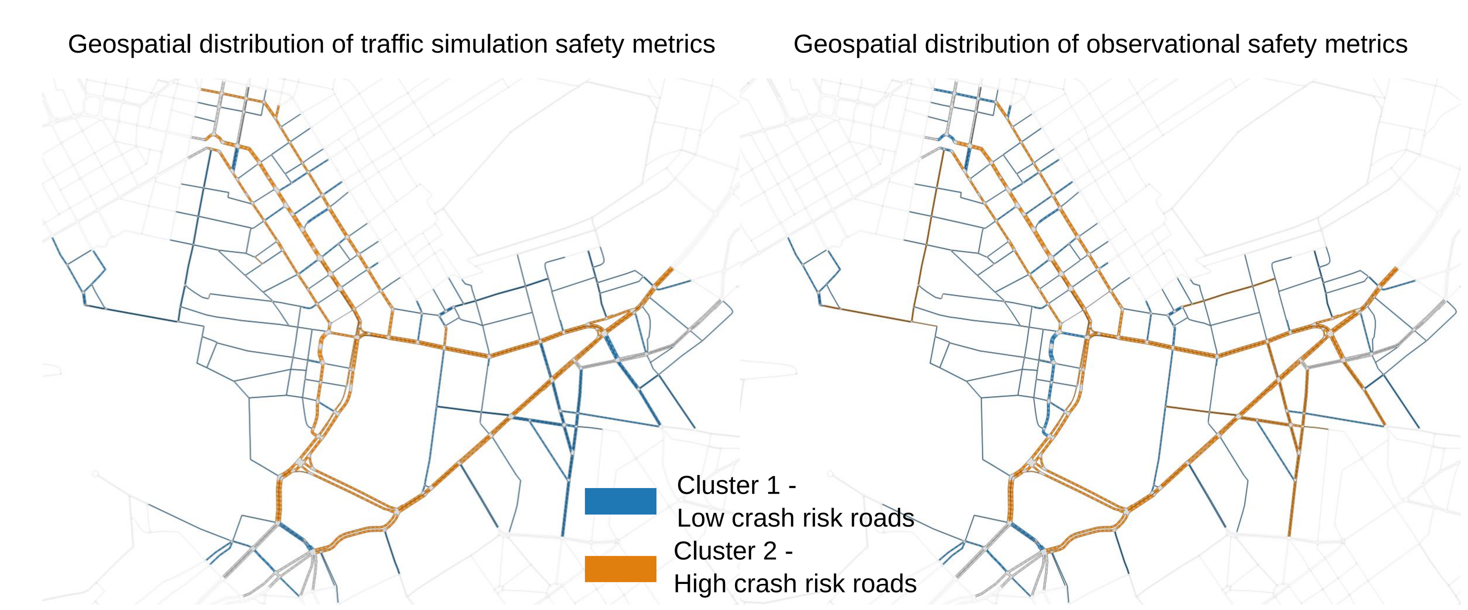
- **Cluster 1:** Roads can be considered as **"low crash risk" roads**, indicating that under both simulated and observed conditions, they exhibit relatively safe traffic dynamics with minimal likelihood of crashes.
- **Cluster 2:** Roads can be considered as **"high crash risk" roads**, indicating that under both simulated and observed conditions, they exhibit significant traffic dynamics with a higher likelihood of crashes compared to low-risk roads.



Data	Cluster	Size	Flow (veh/h)	Capacity (veh/h)	Crash risk	Speed limit (km/h)	Length (m)	Severely injured	Involved vehicles	Crashes
Simulation	1	106	175	546	0.005	-	-	-	-	-
	2	16	1,015	1,898	0.067	-	-	-	-	-
Observed	1	107	-	-	-	26.6	229	0.028	1	1
	2	15	-	-	-	36.8	1,342	0.267	2	19

DISCUSSION

Accuracy was further calculated using a confusion matrix, which yielded **a high accuracy value of approximately 87.7%** (107 out of 122 roads). This indicates that the **traffic simulation model effectively predicts the crash risk classification of roads** when compared to real-world data. This high accuracy validates the simulation method used, demonstrating its reliability and robustness in estimating traffic safety outcomes.



Roads appeared to be irrelevant or inaccurately predicted by the simulation (15 out of 122 roads) correspond to roads that lacked traffic data used for calibration of the simulation model. This finding underscores the **critical importance of including comprehensive and accurate traffic data in the calibration process** of the simulation models. The calibration procedure ensures that the simulation model reflects real-world traffic conditions more accurately, enhancing the reliability of safety predictions as well.

CONCLUSIONS

- By demonstrating traffic simulation accuracy in predicting crash risk classifications and identifying safety patterns, the study establishes **a reliable framework applicable to scenarios where direct observational data is limited or unavailable**.
- The study highlights the importance of thorough calibration; **roads inaccurately predicted lacked sufficient traffic data**, underscoring the need for robust calibration to enhance safety assessment.
- As automation continues to evolve, the methodologies validated in this study will be invaluable in **shaping safer transportation systems** and guiding regulatory efforts to optimize safety standards in increasingly complex traffic environments.

CONVERTING CONFLICTS TO CRASH RISK

A prior study proposed a framework for **calculating crash risk** for every conflict occurred in the simulation network that was also applied in present study:

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

Where:

TTC_i is the Time-To-Collision value for the i^{th} observed conflict and TTC_t is the Time-To-Collision threshold value of the following-vehicle in the i^{th} observed conflict

In present study, the risk assessment took also into account the frequency of traffic conflicts observed on the road that ensured the comparability across different roads:

$$\text{Weighted crash risk} = \frac{\text{Crash risk} \times \text{Number of conflicts}}{\text{Maximum number of conflicts observed}}$$

CLUSTERING METHOD

To validate the followed methodological framework for deriving safety metrics from traffic simulation, **k-means clustering** is employed to differentiate crucial trends and patterns. Specifically, clustering was utilized on both simulated and observed safety metrics **to identify patterns indicative of safety risks** within the study area roads. The evaluation of the cluster centroids describing these categories provides insights into whether traffic behavior differs systematically across various road segments, and whether the simulation accurately reflects real-world conditions.

ACKNOWLEDGMENTS

The present research was carried out within the research project "PHOEBE - Predictive Approaches for Safer Urban Environment", which has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101076963.

Contact

Maria G. Oikonomou
Researcher & Ph.D. Candidate
Department of Transportation Planning and Engineering
National Technical University of Athens
Email: moikonomou@mail.ntua.gr