Identifying Dangerous Street Segments and Analyzing Traffic Behavior in Athens Using Telematics, Crash Records, and TomTom Traffic Data

Aristotelis Tsoutsanis*, Stelios Peithis, Apostolos Ziakopoulos, George Yannis

Department of Transportation Planning and Engineering, National Technical University of Athens (NTUA), Athens, Greece

*Corresponding author: a_tsoutsanis@mail.ntua.gr

Abstract

Traffic safety in Athens, Greece, remains a major concern, particularly on segments where aggressive driving may elevate crash risk. This study introduces a data-driven approach combining telematics and crash records to identify high-risk urban corridors. Telematics data on harsh braking and acceleration inform a Harsh Event Ratio (HER), used to adjust crash counts and derive an Adjusted Crash Score (ACS) for fairer segment comparisons. The top 20 segments by ACS are identified, with eight segments analyzed further using TomTom Traffic API data—average speed, travel time, and delay—to contextualize crash risk. Temporal patterns reveal peaks in harsh driving during morning and evening rush hours, aligning with lower average speeds. While causality is not established due to sample limitations, the integration of behavior, crash, and performance data highlights the role of congestion in driving behavior. The methodology offers a scalable framework for cities with limited volume data to support targeted traffic safety interventions.

Keywords: *traffic safety, Athens, aggressive driving, telematics data, crash risk.*

1. Introduction

Road crash injuries remain a critical global public health issue, exacting a heavy toll in terms of human life, economic cost, and societal impact. According to the World Health Organization (WHO), approximately 1.19 million people die each year due to road traffic crashes, making them the leading cause of death for children and young adults aged 5–29 years (WHO 2023). Despite the existence of proven interventions, progress in reducing fatalities has been insufficient, underscoring the urgent need for comprehensive and sustained efforts to enhance road safety.

In the European context, Greece has consistently exhibited one of the highest rates of road fatalities in urban areas among EU countries. Between 2012 and 2016, Greece suffered over 800 road traffic deaths annually, holding one of the worst performances in the European Union for RTI-related deaths per population (Anagnostou & Cole, 2021). This persistent challenge highlights the necessity for targeted strategies to address the underlying factors contributing to road traffic incidents.

Urban road segments, particularly in densely populated cities like Athens, present unique challenges for traffic safety. Aggressive driving behaviors, such as harsh braking and rapid acceleration, are prevalent on these segments and have been linked to increased crash risks (Ivers et al., 2009). Understanding the dynamics of these behaviors on specific road segments is crucial for developing targeted interventions.

Advancements in telematics and data analytics offer new avenues for understanding and mitigating these risks. By analyzing telematics data, researchers can identify patterns of aggressive driving and correlate them with crash data to pinpoint high-risk road segments. For instance, studies utilizing telematics data have provided insights into driver behavior and its impact on road safety (Ziakopoulos et al., 2022). Moreover, deep learning approaches have been employed to detect aggressive driving behaviors using smartphone GPS sensors, demonstrating high accuracy in identifying such patterns (Talebloo et al., 2021).

Previous research has examined the complex relationship between traffic conditions—such as density, flow, and congestion—and road safety outcomes. Notably, several studies have found that crash rates do not increase linearly with traffic volume. Instead, U-shaped or inverse U-shaped relationships are often observed, where crash risk is elevated during both low and high traffic conditions depending on crash type and flow characteristics (Zhou & Sisiopiku, 1997). Furthermore, traffic congestion has been shown to have mixed effects on road safety: while some studies suggest reduced severity during peak congestion due to lower speeds (Shefer & Rietveld, 1997), others report increased crash rates in heavily congested environments (Kononov et al., 2008). These findings underscore the importance of context-specific and data-driven analyses to better understand how varying traffic dynamics influence driving behavior and safety outcomes.

To address the limitations of relying solely on crash data—often sparse, delayed, or lacking in behavioral context—researchers have increasingly turned to Surrogate Safety Measures (SSMs). These measures, which include indicators such as hard braking, rapid acceleration, and close following distances, offer a proactive means of assessing road safety by capturing high-risk driving behaviors before crashes occur. Telematics, a sub-category of SSMs, provides rich, high-frequency data on vehicle movements and driver behavior, making it a powerful tool for crash prediction and risk assessment. By leveraging SSMs, especially telematics, researchers and policymakers can gain a more immediate and nuanced understanding of safety conditions on the road, enabling more timely and targeted interventions.

This study aims to integrate telematics data, official crash records, and traffic analytics to identify and analyze dangerous road segments in Athens. These behaviors are linked to aggregated crash counts from official records over a six-year period, and an Adjusted Crash Score (ACS) is computed to normalize crash risk across segments. For a subset of locations with complete data, traffic performance metrics such as average speed, travel time, and delay—sourced from the TomTom Traffic API—are incorporated to contextualize risk under varying traffic conditions. By employing regression analysis, the research seeks to establish correlations between aggressive driving behaviors and crash occurrences. The findings are intended to inform data-driven strategies for enhancing road safety and guiding urban policymakers in implementing effective countermeasures.

2. Methodology

This study adopts a multi-source, data-driven methodology aimed at identifying crash-prone street segments in Athens, Greece, and analyzing how driver behavior and traffic flow conditions correlate with elevated crash risk. By integrating official crash statistics, telematics-based behavioral insights, and traffic performance data, a unified framework is established for exploring urban road safety from multiple operational perspectives. The methodology is organized around three primary datasets: (1) historical crash records, (2) mobile telematics data, and (3) traffic analytics sourced via the TomTom Traffic API.

2.1 Crash Data and Spatial Aggregation

The core crash dataset was obtained from local governmental authorities, especially it was captured by the police and maintained by ELSTAT, it consists of traffic collision records spanning a six-year period (2017–2022). A framework that leverages OpenStreetMap (OpenStreetMap contributors, 2017) was developed to match the text of each Street name with the correct coordinates. A graph of edges and nodes of the broader Athens region was created, as shown in Figure 1, from which the coordinates of the streets that had registered crashes were filtered. Each record corresponds to a unique street segment and contains the cumulative number of reported crashes within that segment over the designated period. Due to the absence of point-level geospatial coordinates for individual crash incidents, records were georeferenced using standardized street identifiers.

To spatially contextualize the crash data, a Geographic Information System (GIS)-based spatial join was employed. Specifically, street-level crash counts were aggregated and then joined with known intersection geometries derived from a high-resolution road network shapefile. This allowed for the spatial frequency analysis of crashes, enabling the identification of street segments—with significantly elevated crash densities. These spatially enriched crash metrics serve as the foundation for subsequent cross-dataset integration and prioritization.

Figure 1: *A graph of nodes and edges in the Athens metropolitan area from OpenStreetMap*

2.2 Telematics Data and Behavioral Metrics

The second dataset originates from OSeven Telematics (www.oseven.io), a provider of smartphone-based telematics solutions. This dataset encompasses a broad spectrum of driving behavior events captured during actual trips conducted within the Athens metropolitan area. Each record in the dataset includes a GPS coordinate, a timestamp, and one or more behavioral event flags, with the following key indicators captured:

- Harsh Acceleration: sudden and forceful acceleration likely indicative of aggressive driving;
- Harsh Braking: abrupt deceleration events that may reflect unsafe following distances or inattentive driving;
- Speeding: detected when the driver exceeds the posted speed limit;
- Mobile Phone Usage: captured when a phone is used while the vehicle is in motion.

The dataset includes 25,621 such events collectively, recorded across 6,763 distinct trips. These events are geolocated and timestamped, allowing for precise spatial and temporal aggregation across the urban road network of Athens. Each event was geospatially mapped and then aggregated to the corresponding street segment. Due to spatial and temporal sparsity in the speeding and phone usage data—particularly in low-traffic or peripheral areas—only harsh driving events (acceleration and braking) were retained for use in the core statistical models. Figure 2 presents the spatial coverage of the study, highlighting the street segments covered by the aggregated crash and telematics data.

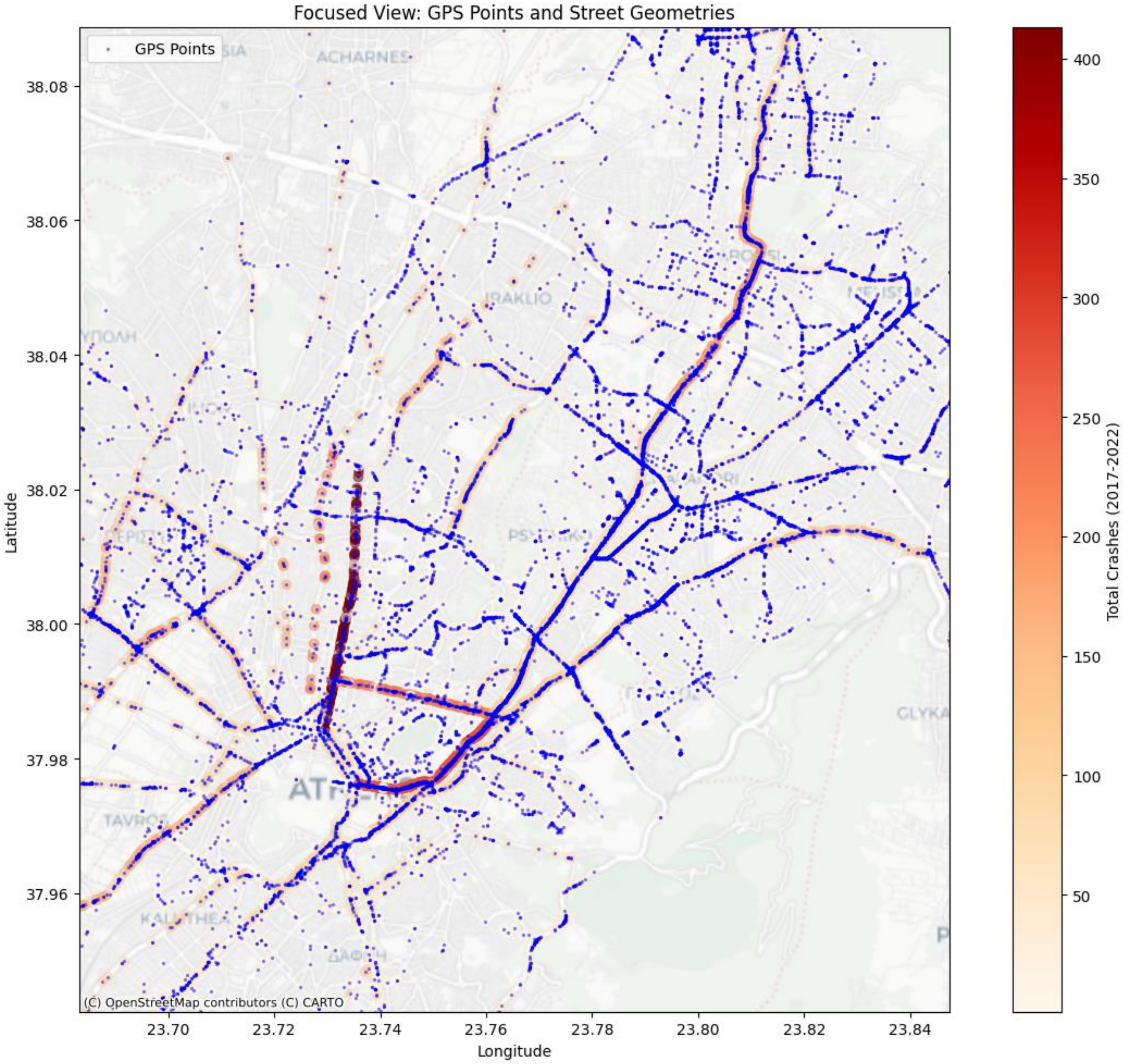

Figure 2: *Subset of the Athens metropolitan area used in the analysis, showing the street segments included in the crash and telematics dataset*

To normalize for differences in data coverage and frequency across the road network, a Harsh Event Ratio (HER) was calculated for each street segment:

$$HER_i = \frac{H_i}{P_i} \qquad \textbf{(1)}$$

where H_i denotes the number of harsh events recorded on street segment i, and P_i is the total number of telematics observation points (GPS pings) recorded for that segment. This ratio represents the proportion of data points characterized by aggressive driving behavior and serves as a standardized measure of behavioral risk. By definition, only segments with telematics data were considered for the analysis,

2.3 Adjusted Crash Scoring for Behavior-Risk Normalization

In the absence of actual traffic volume, raw crash frequencies may not provide a complete picture of risk. To account for this, we introduce an Adjusted Crash Score (ACS) that combines crash frequency with behavioral risk, operationalized through the Harsh Event Ratio:

$$ACS_i = C_i \times HER_i \qquad \textbf{(2)}$$

where C_i is the total number of crashes on segment i over the six-year period from 2017 to 2022, and HER_i is the Harsh Event Ratio for that segment. This derived metric is used to approximate the intensity of crash risk at a location, incorporating both historical crash occurrence and inferred behavioral tendencies from telematics data.

The ACS metric is then used to rank locations and support prioritization in subsequent analyses. This approach allows us to compensate, at least partially, for the lack of direct exposure measures by incorporating behavioral proxies.

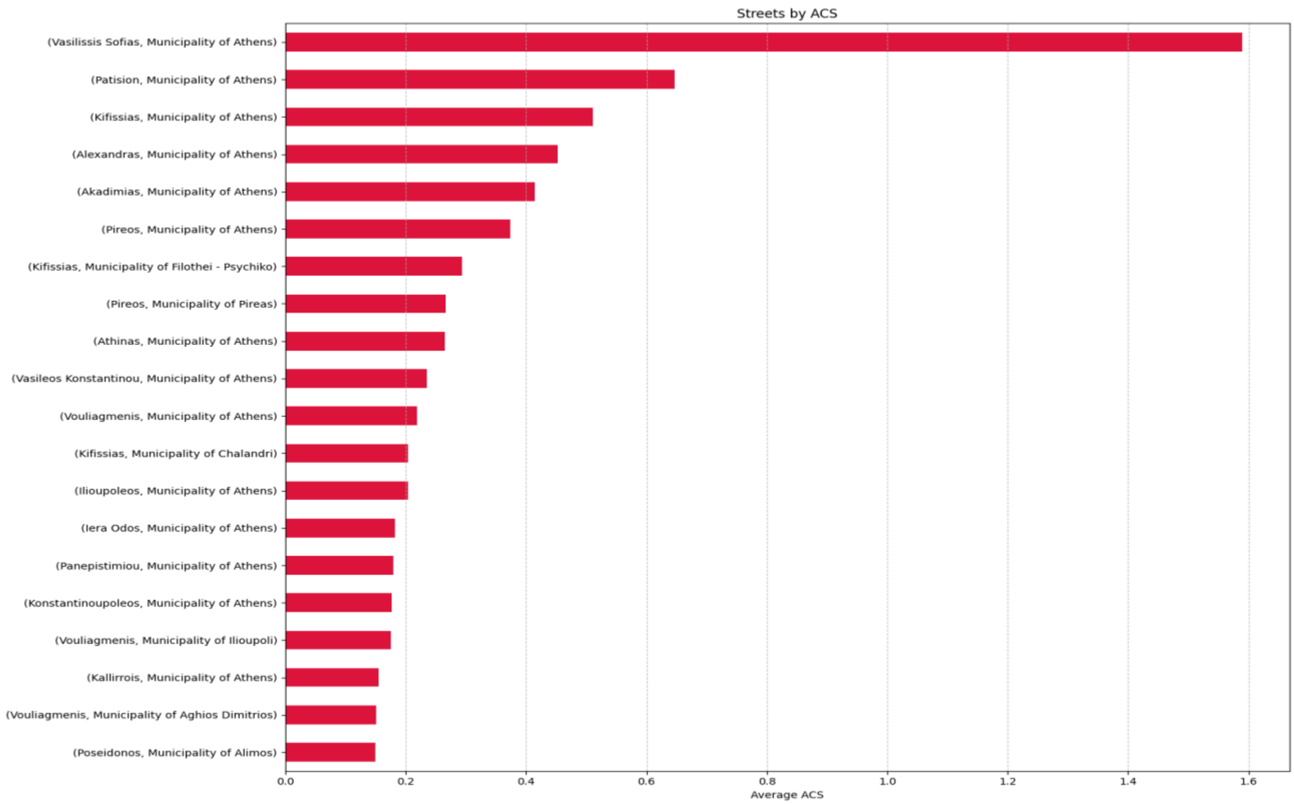

Figure 3: *Bar chart ranking the top 20 most dangerous street segments based on the Adjusted Crash Score (ACS).*

Based on the ACS metric, Figure 3 ranks the most high-risk road segments in the Athens metropolitan area providing a composite measure that reflects both the frequency and severity of traffic incidents. Each bar represents the average ACS for a street segment, aggregated by street name and municipality. Higher ACS values indicate segments with a greater combined risk of frequent and severe events.

2.4 Traffic Flow Data and Integration with Risky Hotspots

The third component of the study involves historical traffic data obtained from the TomTom Traffic API, covering a recent three-month observation period. This dataset contains traffic flow statistics for a curated set of predefined urban routes.

These routes were selected to capture representative traffic patterns across different zones of Athens, including both high-traffic corridors and secondary roads. This enhances the spatial relevance and contextual understanding of traffic conditions in relation to safety indicators.

For each route segment, the traffic dataset provides the following variables at regular collection intervals throughout the day:

- Average Speed (km/h)
- Traffic Delay (seconds)
- Travel Time (seconds)

The resulting dataset provides a continuous profile of traffic efficiency and congestion levels at each location, as summarized in the descriptive statistics presented in Table 1.

Table 1: *Descriptives of traffic data*

name	Travel Time(s)				Traffic Delay(s)			Average Speed(km/h)				Length(m)
	min	mean	max	std	mean	max	std	min	mean	max	std	-
Athinon	51.0	96.7	639.0	45.52	11.19	515.0	42.36	5.55	37.69	62.89	10.85	900.0
Akadimias	175.0	350.41	1423.0	135.15	54.53	729.0	105.28	4.98	15.58	27.67	4.83	1339.0
Alexandras	59.0	138.65	1005.0	78.76	11.85	853.0	43.94	3.51	18.55	32.46	3.7	660.0
Vas. Sofias	226.0	431.72	1249.0	149.36	54.85	747.0	101.31	5.21	14.03	24.12	4.12	1526.0
Vouliagmenis	54.0	86.93	701.0	28.46	1.97	612.0	21.49	4.53	39.1	60.27	8.07	895.0
Ilioupoleos	86.0	170.66	770.0	86.13	23.09	452.0	63.27	5.05	20.24	34.53	5.48	841.0
Kifisias	76.0	176.94	747.0	66.99	40.28	643.0	67.54	4.0	24.38	60.89	10.28	1055.0
Panepistimiou	106.0	237.69	1469.0	139.95	54.34	1104.0	110.62	3.02	18.71	34.91	6.26	1053.0

2.5 Sample Selection

Despite the broader identification of high-risk locations across the Athens metropolitan area, data availability constraints necessitated a focused analytical subset. Specifically, while the ACS was calculated to define the top 20 most hazardous street segments based on integrated crash and behavioral data, 8 out of these 20 segments were also covered by the traffic dataset. This discrepancy arises from the limited geographic overlap between the telematics-crash integration and the predefined route coverage provided by the traffic data. As such, the final stage of analysis focuses on the eight street segments listed in Table 2, which reports the number of harsh events per segment. These segments, shown in Figure 4, represent the subset of high-risk locations with complete crash, behavioral, and traffic flow data available. However, it is important to emphasize that this final stage is underpinned by a robust volume of behavioral data, including 2,835 recorded harsh driving events across these 8 segments. Moreover, the traffic data span a continuous three-month observation period, providing sufficient temporal depth to capture recurring congestion patterns. While this restriction narrows the analytical scope, it ensures methodological consistency and enables a reliable investigation into the relationship between traffic conditions and crash risk.

Table 2: *Number of Harsh Events by Street Segment and Municipality*

Street Name	Harsh Event Count
Athinon – Municipality of Athens	81
Akadimias - Municipality of Athens	111
Alexandras - Municipality of Athens	247
Andrea Syggrou - Municipality of Athens	111
Vasilissis Sofias - Municipality of Athens	1345
Vouliagmenis - Municipality of Ilioupoli	42
Ilioupoleos - Municipality of Athens	34
Kifisias - Municipality of Chalandri	806
Panepistimiou - Municipality of Athens	58

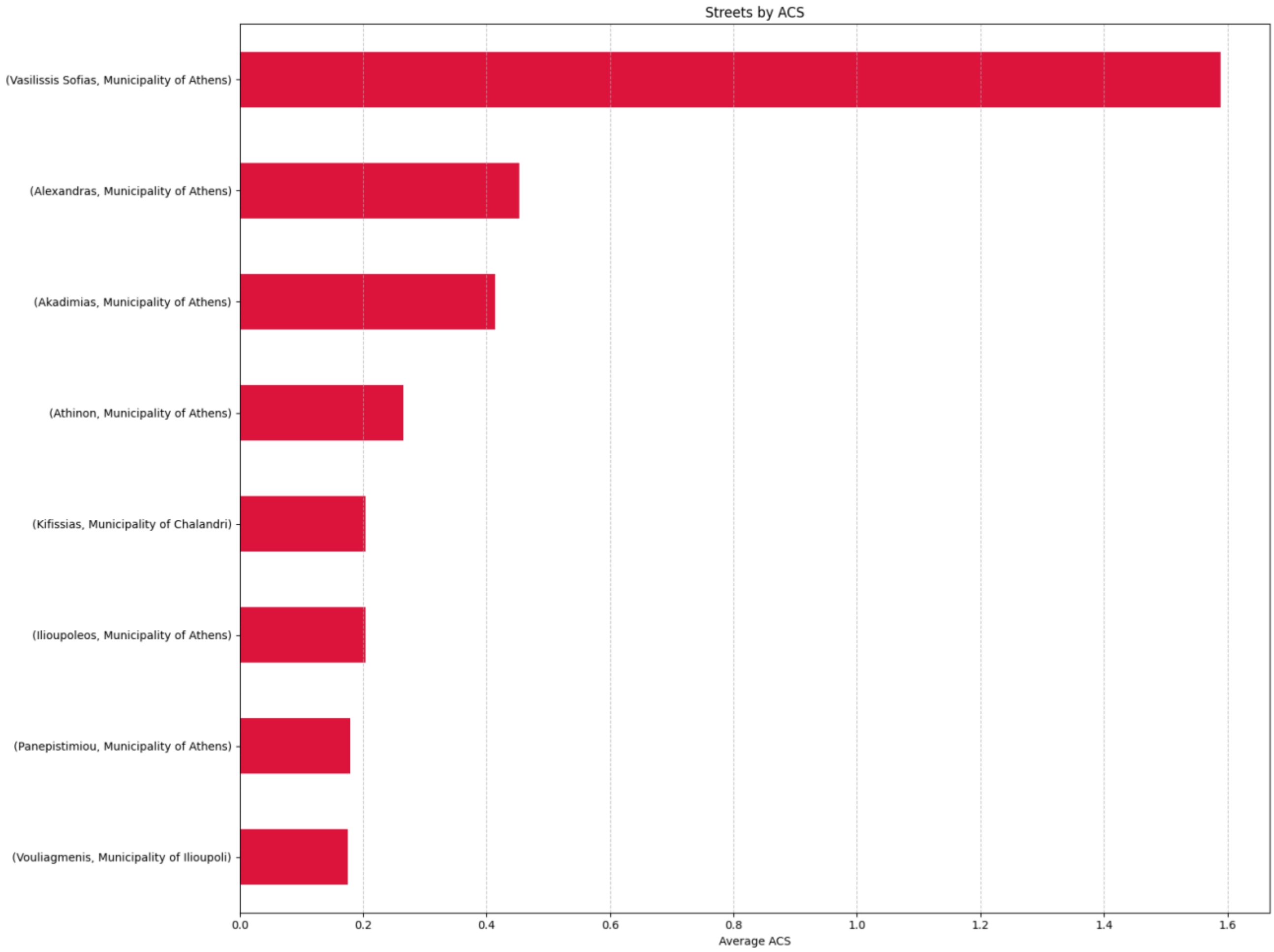

Figure 4: *Bar chart ranking the 8 high-risk street segments by ACS. This figure displays the subset of 8 street segments.*

2.6 Temporal Patterns in Driver Behavior

In addition to aggregate segment-level analysis, we investigated how aggressive driving behavior—proxied by the Harsh Event Ratio (HER)—varies by time of day across the selected segments. As shown in Figure 5, the HER follows a pattern, peaking during the morning (08:00–10:00) and afternoon (17:00–

18:00) rush hours. This pattern reflects elevated behavioral stress during periods of increased congestion, driver interaction, and navigational complexity.

The elevated HER during these periods may not solely reflect reckless driving but rather the increased need for defensive or reactive maneuvers in densely populated traffic conditions. In such settings, frequent braking and abrupt accelerations may be a consequence of maintaining safety margins rather than indicating unsafe intent. This underscores the importance of interpreting telematics behavior within the context of traffic flow and environmental complexity.

It is also important to note that several cells in the heatmap are blank, indicating hours for which no telematics data were available on the respective segments.

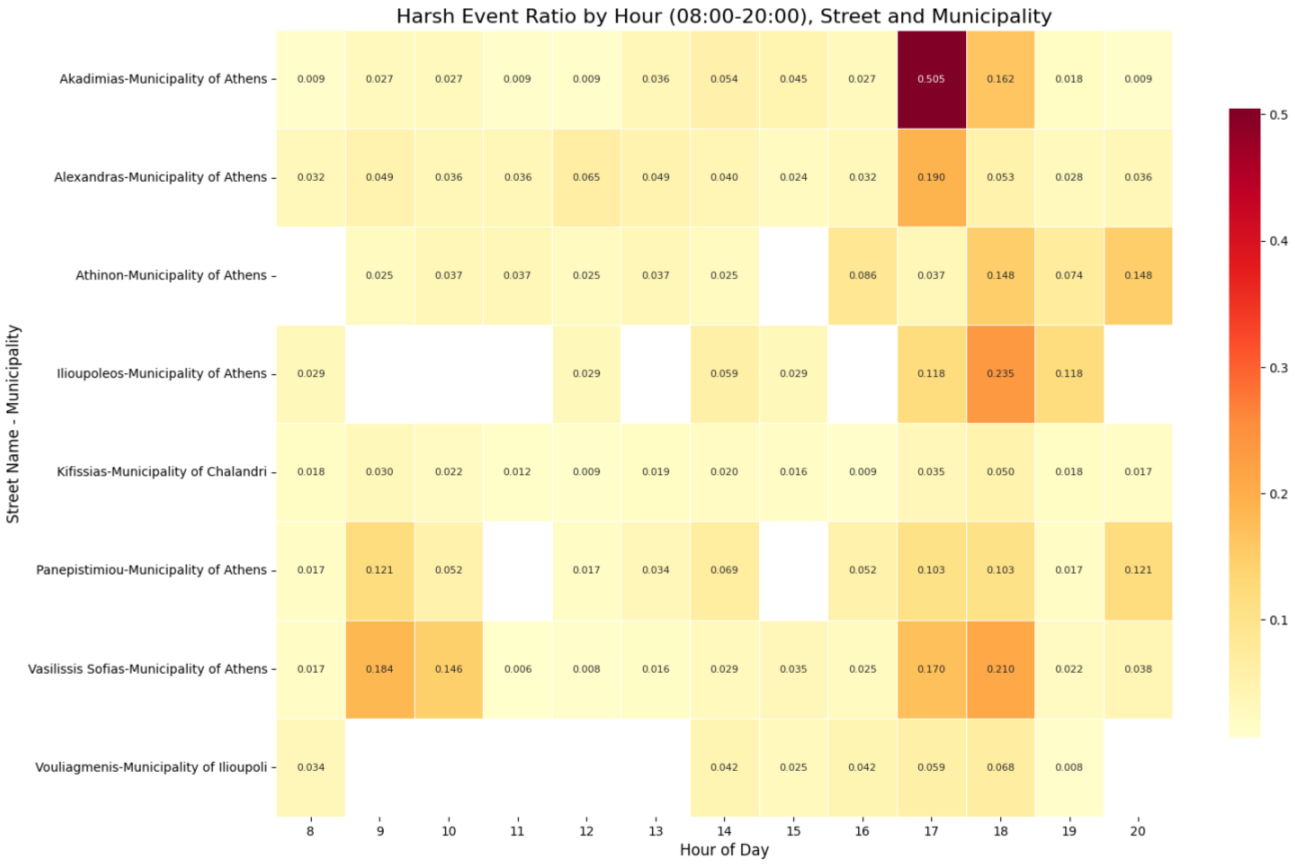

Figure 5: *Hourly variation in Harsh Event Ratio (HER) across the eight selected street segments.*

These periods of heightened behavioral intensity align with the rush hours, as independently validated by Figure 6, which plots the average vehicular speed by hour for the same street segments. The average speed consistently drops during these time windows across most segments, indicating increased traffic density and congestion.

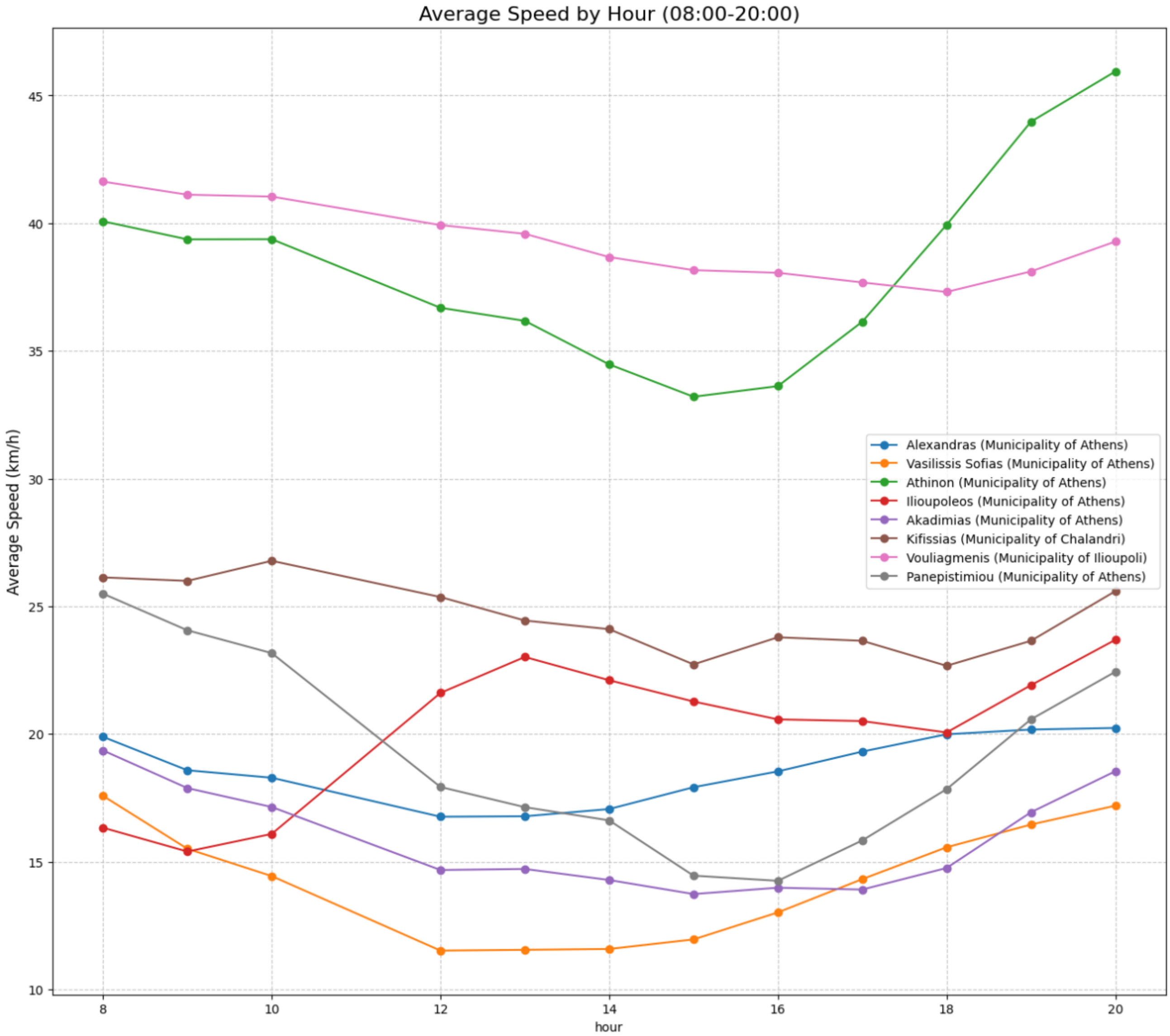

Figure 6: *Hourly variation in average vehicular speed across the eight selected street segments.*

2.7 Average Speed and Adjusted Crash Score

To explore potential trends between traffic performance and crash risk, Figure 7 presents a scatter plot of average speed against the Adjusted Crash Score (ACS) for the eight selected street segments. While a general visual trend suggests that segments with lower average speeds may exhibit higher ACS values, the small sample size of street segments (n = 8) and the heterogeneity of urban traffic conditions do not support definitive statistical conclusions.

Importantly, these results do not imply a direct negative correlation between speed and safety. In high-density urban environments like Athens, lower average speeds often coincide with areas of high conflict density—such as signalized intersections, pedestrian crossings, or mixed-use roadways—where crashes may occur despite slower speeds. Therefore, context-specific factors should be considered when interpreting these results.

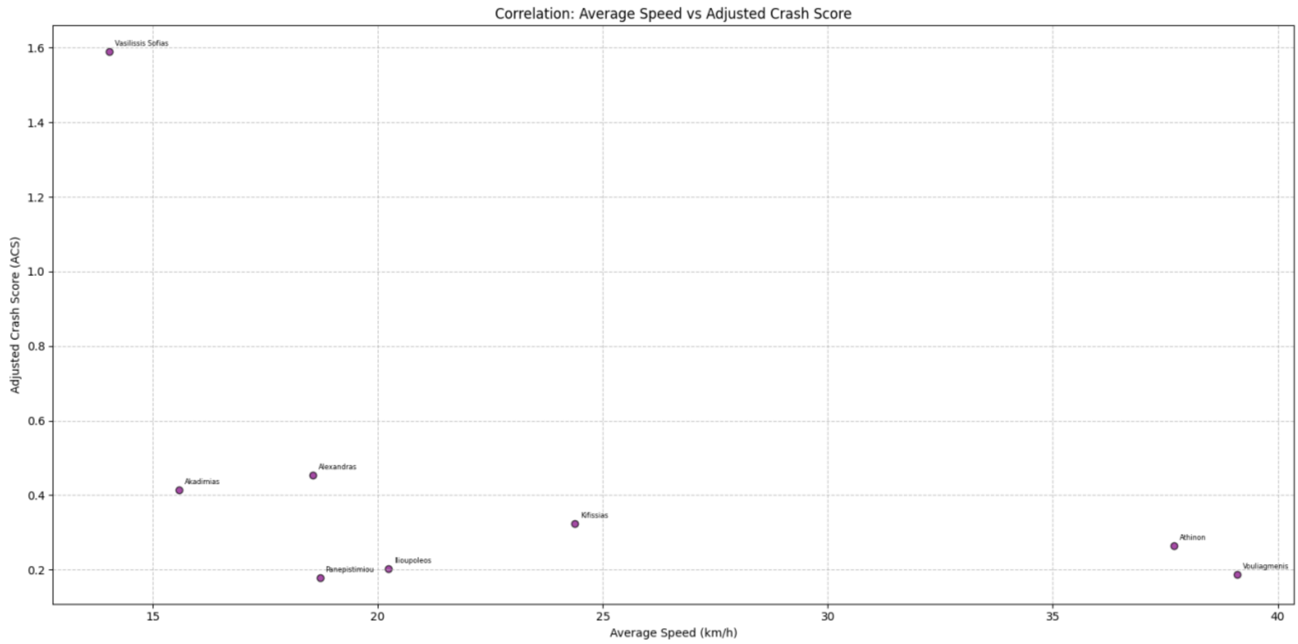

***Figure 7:** Relationship between average speed and Adjusted Crash Score (ACS) across eight high-risk street segments in Athens.*

3. Discussion

The integration of telematics, crash data, and traffic performance analytics offers a robust framework for identifying crash-prone locations and understanding behavioral risk. The introduction of the HER and ACS allow a nuanced, behavior-informed proxy of risk exposure.

While the sample analyzed in the final stage includes eight street segments with full data coverage. These segments are supported by thousands of telematics pings and over 2800 recorded harsh events, enabling meaningful statistical interpretation and offering a representative behavioral snapshot of urban driving patterns.

The observed peak in harsh driving behaviors during rush hours suggests that congestion not only affects flow but also influences driver behavior. However, it is important to interpret harsh events within context: such maneuvers in dense traffic may reflect adaptive or defensive driving rather than inherently unsafe practices.

Limitations of this study include the reliance on aggregated crash data without injury severity information, the lack of direct traffic volume counts, and limited geographic overlap with traffic performance datasets. Despite these limitations, the findings underline the potential of telematics and traffic analytics to inform proactive and location-specific traffic safety strategies, especially in data-scarce urban settings.

4. Conclusion

This study leveraged a multi-source data integration framework to explore the interplay between crash risk, aggressive driving behavior, and traffic performance across selected high-risk street segments in the Athens metropolitan area. By combining official crash records, smartphone-based telematics data, and real-time traffic information from the TomTom API, we were able to derive a behavioral and infrastructural profile of urban streets with elevated crash potential.

The introduction of the Adjusted Crash Score (ACS) enabled a more balanced assessment of crash frequency, taking into account surrogate indicators of exposure through telematics data. This was particularly important given the lack of volume data in the official crash dataset. Through this adjustment, the study identified the top 20 most crash-prone street segments, eight of which had full data coverage for traffic and behavioral metrics and were thus selected for detailed analysis.

Temporal analysis of Harsh Event Ratios (HER) across these segments revealed that aggressive driving behaviors—especially harsh braking and acceleration—peak during morning and evening rush hours. These behavioral patterns coincide with traffic congestion periods as validated through average speed profiles, highlighting the influence of environmental stressors on driver behavior. Importantly, the study emphasizes that harsh events in dense traffic contexts may not solely represent unsafe conduct but can also arise from defensive driving in complex urban conditions.

Nonetheless, this work demonstrates the potential of integrating telematics and traffic data to enhance traditional crash analysis. Such approaches can support data-driven prioritization of safety interventions, especially in dense urban environments where conventional data sources may lack granularity or timeliness.

5. *Acknowledgements*

This research has been conducted within the IVORY project. The project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101119590.

6. *References-Bibliography*

Anagnostou, E., & Cole, E. (2021). Targeted prevention of road traffic deaths in Greece: A multifactorial 5-year census-based study. European Journal of Trauma and Emergency Surgery, 47(4), 1137–1152. https://doi.org/10.1007/s00068-019-01290-3

OpenStreetMap contributors. (2017). Planet dump retrieved from https://planet.osm.org.

Ivers, R., Senserrick, T., Boufous, S., Stevenson, M., Chen, H.-Y., Woodward, M., & Norton, R. (2009). Novice drivers' risky driving behavior, risk perception, and crash risk: Findings from the DRIVE study. Am. J. Public Health, 99(9), 1638–1644.

Kononov, J., Bailey, B., & Allery, B. (2008). Relationships Between Safety and Both Congestion and Number of Lanes on Urban Freeways. Transportation Research Record, 2083, 26–39. https://doi.org/10.3141/2083-04

Shefer, D., & Rietveld, P. (1997). Congestion and Safety on Highways: Towards an Analytical Model. Urban Studies, 34(4), 679–692. https://doi.org/10.1080/0042098975970

Talebloo, F., Mohammed, E. A., & Far, B. H. (2021). Deep Learning Approach for Aggressive Driving Behaviour Detection. CoRR, abs/2111.04794. https://arxiv.org/abs/2111.04794

World Health Organization, 2023. Global status report on road safety, https://iris.who.int/bitstream/handle/10665/375016/9789240086517-eng.pdf?sequence=1

Zhou, M., & Sisiopiku, V. P. (1997). Relationship Between Volume-to-Capacity Ratios and Rates. Transportation Research Record, 1581(1), 47–52. https://doi.org/10.3141/1581-06

Ziakopoulos, A., Vlahogianni, E., Antoniou, C., & Yannis, G. (2022). Spatial predictions of harsh driving events using statistical and machine learning methods. Safety Science, 150, 105722. https://doi.org/10.1016/j.ssci.2022.105722