# Exploring the Relationship Between Unsafe Traffic Events and Crash Occurrences Using Smartphone App Data

Paraskevi Koliou[1a], Apostolos Ziakopoulos[1], Virginia Petraki[1], George Yannis[1]
[1]Department of Transportation Planning and Engineering/School of Civil Engineering, National Technical University of Athens/NTUA, Athens, Greece [a] Corresponding author

## Abstract

Crash occurrence analysis is a traditional method for assessing traffic safety, yet more accurate or evident crash records may be necessary. However, unsafe traffic events such as harsh acceleration/braking instances occur more frequently and can be readily obtained. This study investigates the relationship between crash frequency and the occurrence of unsafe traffic events - harsh acceleration and braking events - utilizing smartphone app data across a network. The research aims to assess whether crashes can be predicted based on this data. Acceleration/braking events will be extracted from smartphone app data, enabling an analysis of their spatiotemporal distribution. This research explores whether the type of traffic events and their spatiotemporal resolution can enhance the prediction of crashes at specific sites such as intersections. Various regression models are developed and evaluated to determine the most accurate and reliable crash prediction models based on the combination of unsafe traffic events and spatiotemporal resolution. The anticipated findings will advocate for proactive approaches to traffic safety analysis and delineate the minimum requirements of unsafe traffic event data for such analysis.

**Keywords:** traffic road safety analysis, unsafe traffic events, crash prediction, big data application, generalized linear models,

## 1. Introduction

### 1.1 Background and Motivation

Traffic safety is an important concern for urban planners, transportation engineers, and public safety officials. Traditional methods relying on crash data have limitations due to the sporadic nature and underreporting of crashes. There is growing interest in using more frequent indicators like harsh acceleration/braking events, which occur more often and can be tracked via smartphone technology. This study leverages high-resolution data from smartphone sensors, combined with traffic metrics and road characteristics, to investigate driver behavior at intersections in Athens, specifically Mesogeion and Vouliagmenis avenues.

Using GIS for spatial mapping and developing Generalized Linear Models (GLM), the research reveals that increased traffic load per lane correlates with more abrupt events, while higher average occupancy at intersections leads to frequent sudden accelerations and higher traffic speeds result in sudden decelerations. This highlights the significance of traffic metrics over road characteristics in influencing abrupt driving events. The integration of smartphone sensor data with GIS allows for detailed spatial analysis, identifying areas needing intervention to improve traffic management and safety. The study's findings suggest that unsafe traffic events can predict crash occurrences. By analyzing the spatiotemporal distribution of these events and developing regression models, the research aims to improve crash prediction accuracy at specific sites, such as intersections.

### 1.2 Road Safety State

Despite ongoing efforts to reduce road crashes and fatalities, global statistics have plateaued. In 2018, road crashes caused 1.35 million deaths annually, or about 3,700 fatalities per day worldwide (WHO, 2018). In the European Union, there were approximately 20,653 road fatalities in 2022, a 4% increase from 2021, yet a 10% decrease from 2019 (European Commission, 2023).

Fatality rates vary significantly across Europe. Sweden and Denmark report the lowest rates, with 22 and 26 deaths per million inhabitants respectively, while Romania and Bulgaria have the highest rates, with 86 and 78 deaths per million inhabitants respectively (Eurostat, 2023). Greece achieved a 51% reduction in crash fatalities between 2009 and 2018 but still ranks 22nd among EU states, with 58 deaths per million inhabitants in 2022, slightly up from previous years (European Commission, 2019; Eurostat, 2023). Economic recession has been partially credited for the reduction in fatalities (Yannis et al., 2014). However, the Hellenic Statistical Authority (ELSTAT, 2020) reported an 18.8% increase in road crashes causing death or injury in January 2018 compared to January 2017 (European Commission, 2023).

Overall, while progress has been made in some areas, the overall reduction in road fatalities across Europe remains slow, with significant disparities between different countries. The EU aims to halve road deaths by 2030 as part of its Vision Zero strategy, but reaching this target will require sustained and coordinated efforts across all member states (European Commission, 2024).

## 1.3 State of the Art

### 1.3.1 The Importance of Driving Behavior Analysis

Analyzing driver behavior is crucial for preventing road crashes and improving road safety, as human factors contribute to about 95% of road crashes (Singh, 2015). The NHTSA defines the critical reason for a crash as "the immediate reason for the critical pre-crash event and is often the last failure in the causal chain of events leading up to the crash" (Singh, 2018). Driving behavior, influenced by various factors, plays a significant role in road crashes (Dingus et al., 2016), and the road environment also impacts driving behavior (Horberry et al., 2006; Hamdar et al., 2016).

Given current road safety challenges, exploring new approaches to crash reduction is essential, including using smartphone applications for data collection and analysis (Vlahogianni and Barmpounakis, 2017). Smartphones, equipped with sensors such as accelerometers, gyroscopes, and GPS, are valuable tools for transport studies and sensing applications without user engagement (Mantouka et al., 2018).

In the car insurance market, analyzing driving behavior is also valuable. Programs like Pay-As-You-Drive (PAYD) charge drivers based on location and time of driving, promoting responsible driving and potentially reducing crash risk (Troncoso et al., 2010). The insurance industry has explored the correlation between aggressive driving behaviors, such as harsh accelerations and braking, and crash risk (Paefgen et al., 2014; Tselentis et al., 2017).

### 1.3.2 Exploitation of Sensor Data

Analyzing driver behavior is crucial, yet collecting reliable and high-resolution data presents challenges. Data can be gathered through questionnaires, simulators, in-vehicle data recorders, or smartphone sensors. Smartphones offer a low-cost, easy-to-install platform for detecting driver behavior in naturalistic conditions (Papadimitriou et al., 2018). Modern mobile technologies using internal sensors provide real-time feedback on driving behavior, promoting safety and potentially reducing crashes by about 20% under specific conditions (Wouters and Bos, 2000).

Studies have utilized vehicle-integrated systems and smartphone sensors to examine driving behavior, providing high-resolution, objective measurements and enabling real-time feedback to drivers. Feedback has been shown to encourage more careful and responsible driving (Roetting et al., 2003; Toledo et al., 2008). While there are significant correlations between vehicle recording systems and smartphone sensors, factors such as event type, smartphone location in the car, and external conditions can affect data quality (Paefgen et al., 2012). Due to the high cost of in-vehicle systems, smartphone applications offer a feasible alternative for data collection.

Additionally, spatial analysis of recorded road crashes, road characteristics, and census variables, along with geographic information systems, enhances road safety models (Ziakopoulos and Yannis, 2020a; Ziakopoulos and Yannis, 2020b; Abdel-Aty et al., 2013).

### 1.3.3 Merits of Harsh Event Analysis

Analyzing road crashes is traditional in road safety science, but harsh events like accelerations and braking offer additional insights. These events correlate strongly with reduced spatial and temporal headways, near misses, and other risky behaviors (Tselentis et al., 2017; Bonsall et al., 2005; Gündüz et al., 2017; Paefgen et al., 2014).

Harsh events can be analyzed as point-data, similar to crashes, revealing patterns and dependencies with independent parameters. An aggressive driver will exhibit elevated harsh events across all trips, indicating high-risk road segments (hotspots). These events provide proactive safety parameters, identifying hotspots before crashes occur. They are increasingly used in usage-based motor insurance (UBI) to represent crash occurrence probability (Tselentis et al., 2017).

Research on factors influencing harsh events is limited compared to crash analysis, revealing significant gaps. Harsh accelerations and braking occur in different contexts and should not be analyzed collectively. Drivers with higher anger, frustration, and anxiety levels display higher acceleration values (Stephens and Groeger, 2009). Harsh braking events indicate reactions to safety-critical situations and are used as indicators in naturalistic driving studies (Hanowski et al., 2005; Olson et al., 2009; Zohar et al., 2014; Jansen and Wesseling, 2018).

### 1.3.4 Traffic Safety Assessment Methods

Traditional methods of traffic safety assessment have relied heavily on crash data from police reports, hospital records, and insurance claims. While these methods provide direct measures of traffic safety issues, they have significant limitations, including the rarity of crashes and potential underreporting (Hauer, 1997). This makes statistical analysis challenging due to small sample sizes.

Near misses and other unsafe traffic events, such as harsh acceleration and deceleration, occur more frequently than crashes and can serve as early indicators of potential hazards. Research indicates that these events often precede crashes and can identify hazardous locations before crashes occur (Laureshyn, 2010). Intersections with high rates of near misses tend to have higher subsequent crash rates (Archer, 2005), allowing for a more proactive traffic safety strategy.

The advent of smartphones has revolutionized data collection in traffic safety research. Equipped with sensors like accelerometers and gyroscopes, smartphones capture detailed information about driving behaviors, including sudden stops, sharp turns, and rapid accelerations. Studies have shown the potential of smartphone data to provide real-time insights into driver behavior and traffic conditions (Barić et al., 2014). Smartphone applications for driver assistance and monitoring have been used to collect extensive data on unsafe traffic events, which can be analyzed to identify patterns and correlations with crash occurrences (Zhao et al., 2017). This enables continuous and widespread monitoring of traffic conditions, surpassing the capabilities of traditional methods.

While traditional crash data remains vital, its limitations necessitate alternative data sources. Near misses and harsh driving events captured via smartphone apps complement crash data, allowing for more frequent and comprehensive traffic safety monitoring. However, integrating these new data sources into predictive models is an emerging field with several gaps. Key gaps include understanding the relationship between different types of unsafe events and actual crashes, determining the optimal spatiotemporal resolution for analyzing these events, and developing robust predictive models to forecast crashes based on unsafe event frequency and distribution.

Leveraging smartphone app data to predict crashes represents a significant advancement in traffic safety research. This study aims to address some existing gaps by exploring the predictive power of near misses and harsh driving events, contributing to more proactive and effective traffic safety management strategies.

## 2. Methodology and Theoretical Background

### 2.1 Data Exploitation

Modelling driver behavior is a complex phenomenon that has long interested the scientific community. This study aims to investigate the combined influence of road characteristics and traffic on driver behavior, particularly in crash occurrence, using smartphone data on harsh acceleration and braking events in an urban intersection environment. Building on the work of Petraki et al. (2020), the research examines how the road environment and traffic conditions affect driving behavior at intersections, focusing on abrupt accelerations and braking. Conducted at a macroscopic level, the study area includes two major urban expressways in Athens—Mesogeion Avenue and Vouliagmenis Avenue—chosen for their similar traffic lane configurations and separated travel directions. These avenues provide a suitable context for analyzing the impact of road and traffic characteristics on driver behavior (see **Error! Reference source not found.**).

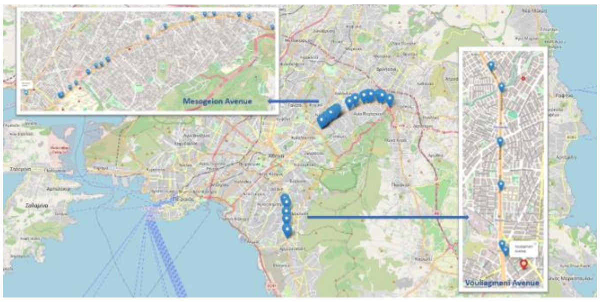


*Figure 1:Research Area - Mesogeion and Vouliagmenis Avenues.*

The data analyzed in this study were sourced from three primary sources. First, driving behavior data were collected from approximately 300 drivers in Athens using the OSeven smartphone application, which records driving behavior. This data captures instances of unsafe traffic events, specifically harsh acceleration (HA) and braking (HB) events. The dataset includes metrics pertinent to traffic safety, such as the identification of junctions where specific events were recorded, traffic volume, average speed, and occupancy rate, providing a comprehensive overview of traffic conditions.

Secondly, traffic metrics were obtained from the Traffic Management Center of the Attica region. These metrics, including traffic volume and average speeds, were collected through 26 loops installed at specific measurement points along the two studied urban expressways. Lastly, road characteristics were extracted using the Google Maps online mapping service, detailing features of road segments and intersections, including lane numbers and configurations. During data collection and processing, challenges were addressed to ensure dataset quality and reliability by standardizing data units and formats in Excel for consistency between sources, and by ensuring accurate spatial alignment in QGIS using precise geolocation data from Google Maps and cross-referencing with known traffic loop locations.

Data integration involved merging driving behavior data from OSeven with traffic metrics from the Traffic Management Center and road characteristics from Google Maps. Spatial mapping using QGIS correlated abrupt driving events with specific road segments and intersections, resulting in a comprehensive database for analyzing harsh acceleration and braking events on the examined avenues. The OSeven application identifies harsh events through data fusion and machine learning algorithms, integrating input from accelerometers, GPS, gyroscopes, and other sensors, rather than using predefined thresholds. This methodology, supported by previous studies (Yannis et al., 2017; Tselentis et al., 2018, 2019; Stavrakaki et al., 2019; Petraki et al., 2020; Papadimitriou et al., 2019), has proven effective in road safety research. The application adheres to GDPR regulations, ensuring no additional user information is collected, and its flexibility facilitates data acquisition without extensive vehicle instrumentation or costly methods (Ziakopoulos et al., 2020).

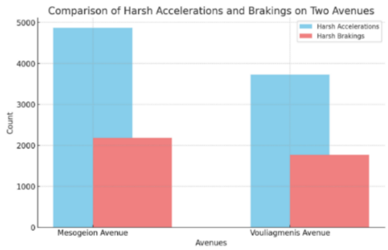


*Figure 2: Comparison of Harsh Accelerations and Braking on the two avenues.*

A total of 303 drivers participated in a naturalistic driving experiment conducted in Athens between August 25, 2016, and November 26, 2017, resulting in the creation of extensive databases of harsh acceleration and deceleration events. Specifically, during this period, 4,869 harsh accelerations and 2,181 harsh braking were recorded on Mesogeion Avenue, while 3,723 harsh accelerations and 1,765 harsh braking were documented on Vouliagmenis Avenue (see **Error! Reference source not found.**).

## 2.2 Statistical Background

The data was analyzed at varying spatial and temporal resolutions to explore the relationship between unsafe traffic events and crash occurrences, focusing on harsh acceleration and braking events. Spatial resolution was assessed at the level of road intersections (Junctions of Mesogeion - JK and Junctions of Vouliagmenis - JV), while temporal resolution was evaluated on a monthly, weekly, and daily basis. GIS tools were used to map each unsafe traffic event to specific sites, facilitating the analysis of their spatial distribution in relation to crash occurrences. The spatial analysis was followed by statistical analysis, expanding on Petraki et al. (2020) by including further investigation into Speed Difference and Event Speed (minimum, maximum, and standard deviation). This investigation aimed to identify high correlations between dependent variables and influencing factors, using the Generalized Linear Model (GLM).

Generalized Linear Models are particularly suitable for transportation research due to their flexibility in handling various types of data distributions, such as binomial, Poisson, and normal distributions. Previous studies have demonstrated the effectiveness of GLMs in modeling relationships where the response variable does not necessarily follow a normal distribution, making them ideal for analyzing transportation data such as accident counts or binary outcomes (accident occurred or not).

The Generalized Linear Model (GLM) uses a linear predictor to model the log odds of the outcome. According to McCulloch (2008), if $y_i$ represents the observed frequency of harsh events per trip i (considering harsh braking and harsh acceleration separately), and $\lambda_i$ represents the expected frequency of these events, the model can be specified as $yi \sim Poisson\ (\lambda_{i)}$ (1). The linear predictor in this case is expressed as in equation (2):

$$log\ (\lambda i) = \beta_0+\beta_n x_n+\varepsilon \quad (2)$$

Here, β includes the fixed-effect parameters (constant and coefficients) for the n independent variables, and ε is the error term. The GLM can be expanded into a Generalized Linear Mixed Model (GLMM) by incorporating random effects. In GLMMs, random effects are represented as random variable coefficients (random slopes) or random intercepts as in Equation (3):

$$log\ (\lambda i) = \beta_{0i} + \beta_{ji}x_{ji} + \beta_{n-1}x_{n-1} + \varepsilon \quad (3)$$

In this formulation, $\beta_{0i}$ and $\beta_{ji}$are assumed to follow normal distributions centered around the values of their corresponding fixed effects:

$$\beta_{0i} \sim N\ (\beta_{0,}\ \sigma^2{}_{s0}) \quad (4)$$

$$\beta_{ji} \sim N\ (\beta_{0,}\ \sigma^2{}_{s0}) \quad (5)$$

Interpreting these coefficients is more straightforward when using relative risk ratios (also known as incidence rate ratios). These ratios transform the predictor to reflect the frequency. For an increase of one unit in a specific variable k, while keeping all other parameters constant, the predicted frequency $\lambda_i$ is scaled by Equation (6):

$$\lambda_{ki} = exp\ (\beta_{ki}) * \lambda_i \quad (6)$$

As noted by McCulloch (2011), models with random effects can accommodate correlated independent variables, thus overcoming some of the limitations of traditional

GLMs. Additionally, to facilitate the fitting of the GLMM, the trip data was standardized using z-score scaling. This is a common practice that does not alter the coefficients obtained from the analysis.

The model that best fits the data—containing the most informative variable combination and explaining the highest degree of variance—is chosen based on the minimum AIC. It's important to emphasize that the contribution of any random effects is evaluated by performing a custom ANOVA (log-likelihood test) comparing the fixed effects GLM with any developed GLMMs.

For both Ha and Hb models, a mixed-effects model with random intercepts (GLMER) was used alongside the base GLM. This included random intercepts for junctions on Mesogeion Avenue (JM) and Vouliagmenis Avenue (JV). A likelihood ratio test confirmed the GLMER model's significant improvement over the GLM. The regression models examined the influence of driving behavior, traffic characteristics, and road geometry on harsh event frequency, with all variables except one being statistically significant. Multicollinearity was checked using VIF values, all below 5, ensuring robustness. This approach highlights the importance of local variations and specific road characteristics in traffic safety analysis, providing a solid foundation for targeted interventions. The following sections present the analysis results and discuss their implications.

## 3. Results and Discussion

### 3.1 Descriptive Statistics of the Data

The dataset presents traffic event data for various junctions, classified into two types, Junctions of Mesogeion (JM) and Junctions of Vouliagmenis (JV). The data includes metrics on vehicle flow (Q[Veh/h]), average speed (V [km/h]), occupancy (O[%]), and various statistics on speed differences and distances. This analysis aims to explore the relationship between unsafe traffic events, such as harsh acceleration/braking, and crash occurrences.

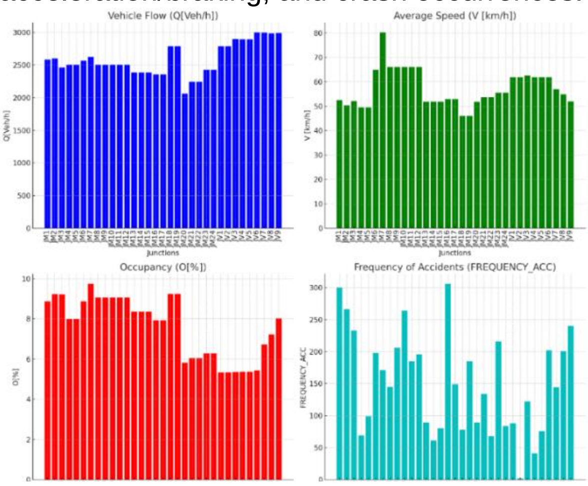

*Figure 3:Visual representation of key traffic metrics*

The dataset includes detailed measurements for each junction, categorized by junction type, with key variables such as the number of lanes, vehicle flow rate, average speed, and metrics on speed differences and distances.  visually represents these traffic metrics, showing variations in vehicle flow, speed, occupancy, and accident frequency across different junctions, identifying this way high-risk junctions, enabling targeted safety and efficiency measures.

Vehicle flow (Q[Veh/h]) ranges from 2061.391 (JM20) to 3001.898, with the highest flow observed at Junction 6 of Vouliagmeni (JV6). The average speed (V [km/h]) varies significantly, with Junction 7 of Mesogeion (JM7) having the highest average speed of 80.237 km/h and Junction 9 of Vouliagmeni (JV9) the lowest at 51.869 km/h. This variation indicates different traffic conditions and congestion levels at various junctions.

Metrics on speed differences such as mean_Speed_Diff, min_Speed_Diff, and max_Speed_Diff provide insights into the variability of speeds at different junctions. For instance, JV9 has the highest max_Speed_Diff of 30.946, indicating significant speed variations which could be a risk factor for crashes. The dataset also includes event-specific speed metrics such as mean_Event_Speed and range_Event_Speed. High values in these metrics, such as the range_Event_Speed of 75.010 at JM15, suggest significant fluctuations during events, potentially indicating harsh braking or acceleration.

Occupancy (O[%]) values highlight the percentage of time the junction is occupied by vehicles. Higher occupancy rates, like the 9.749% at JM7, can correlate with higher traffic density and potential congestion. Frequency_acceleration denotes the frequency of accident occurrences, with values like 306 at JM16 indicating a higher incidence of crashes. Distance metrics such as mean_distance, min_distance, and max_distance provide spatial insights. For example, JM9 has a maximum distance (max_distance) of 152.245, which might indicate larger junctions or intersections, potentially affecting traffic flow and safety.

### 3.2 Statistical Modelling Results

In this section, the outcomes of the statistical analyses are being presented. The GLM analysis involved fitting a Poisson regression model to predict crash frequency, based on various predictors. The summary of the GLM provides key insights into the coefficients and their significance levels (p-values), which indicate the strength and reliability of each predictor. Additionally, model fit statistics such as AIC and Deviance were calculated to assess the model's overall performance. These metrics provide a quantitative measure of how well the model explains the data.

To account for within-group correlation, a Generalized Linear Mixed Model (GLMM) was fitted, including a random intercept for `Junct_Type`. The GLMM summary highlights the coefficients and random effects, along with their significance levels. This model offers an enhanced understanding of the variability within junction types, which is not captured by the GLM. Similar to the GLM, model fit statistics such as AIC and Deviance were reported, allowing for a direct comparison between the two modeling approaches.

To ensure the robustness of the predictors, VIF values were calculated to assess multicollinearity. High VIF values indicate potential redundancy among predictors, which can affect the model's stability and interpretability. By addressing multicollinearity, the reliability of the model coefficients is being improved.

Results of the performed log-likelihood ratio test are being conducted, comparing the GLM and GLMM models. This test evaluates whether the inclusion of random effects significantly improves the model fit, providing statistical justification for the use of GLMM over GLM. Furthermore, a caterpillar plot was created to visualize the random effects from the GLMM model. This plot aids in the interpretation of the random effects associated with different junction types, highlighting variations that may influence crash frequencies. The visual representation provides a clear and intuitive understanding of how junction types contribute to the model.

In summary, the results of the statistical analyses, including GLM and GLMM summaries, VIF values, log-likelihood ratio test, and caterpillar plot, collectively offer a comprehensive understanding of the factors influencing crash frequencies. These findings underscore the importance of considering both fixed and random effects in transportation safety research. To model the expected frequency of accelerating (Ha) and braking (Hb) events, GLM-based models were calibrated as previously explained. Given the high-resolution, big-data collection scheme, random effects were included to capture unique driving behavior traits for each driver.

| Exploring the relationship between unsafe traffic events and crash occurrence using harsh acceleration events | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| glm(formula = FREQUENCY_ACC ~ MIN_Speed_Diff + MAX_Speed_Diff + MAX_Event_Speed + STD_Event_Speed + Right_Exits + Outgoing_Lanes + Sideway, family = poisson, data = Vdata) | | | | | | | | | | | |
| | | | | | | | | | | | |
| Base GLM model results | | | | | | Random intercepts GLMER model: | | | | | |
| Coefficients: | | | | | | Fixed effects: | | | | | |
| | Estimate | Std.Error | z value | Pr(>\|z\|) | | | Estimate | Std. Error | z value | Pr(>\|z\|) | |
| (Intercept) | 3.901915 | 0.135481 | 28.8 | < 2e-16 | *** | (Intercept) | 3.609952 | 0.224784 | 16.06 | < 2e-16 | *** |
| MIN_Speed_Diff | -0.10434 | 0.011044 | -9.448 | < 2e-16 | *** | MIN_Speed_Diff | -0.082331 | 0.011598 | -7.099 | 1.26E-12 | *** |
| MAX_Speed_Diff | 0.0821 | 0.005783 | 14.196 | < 2e-16 | *** | MAX_Speed_Diff | 0.076991 | 0.00591 | 13.027 | < 2e-16 | *** |
| MAX_Event_Speed | 0.018499 | 0.001451 | 12.752 | < 2e-16 | *** | MAX_Event_Speed | 0.015155 | 0.001523 | 9.954 | < 2e-16 | *** |
| STD_Event_Speed | -0.128158 | 0.008623 | -14.863 | < 2e-16 | *** | STD_Event_Speed | -0.108751 | 0.009211 | -11.807 | < 2e-16 | *** |
| Right_Exits | 0.299347 | 0.026149 | 11.448 | < 2e-16 | | Right_Exits | 0.198172 | 0.030031 | 6.599 | 4.14E-11 | *** |
| Outgoing_Lanes | -0.126292 | 0.016905 | -7.471 | 0.00000000000008 | *** | Outgoing_Lanes | 0.003163 | 0.024918 | 0.127 | 0.898988 | |
| Sideway | 0.046146 | 0.03461 | 1.333 | 0.1820 | | Sideway | 0.119132 | 0.035797 | 3.328 | 0.000875 | *** |
| --- | | | | | | | | | | | |
| Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 | | | | | | | | | | | |

| Exploring the relationship between unsafe traffic events and crash occurrence using harsh braking events | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| glm(formula = FREQUENCY_BRK ~ MIN_Speed_Diff + MAX_Speed_Diff + MAX_Event_Speed + STD_Event_Speed + Right_Exits + Outgoing_Lanes + Sideway, family = poisson, data = Vdata) | | | | | | | | | | | |
| | | | | | | | | | | | |
| Base GLM model results | | | | | | Random intercepts GLMER model: | | | | | |
| Coefficients: | | | | | | Fixed effects: | | | | | |
| | Estimate | Std.Error | z value | Pr(>\|z\|) | | | Estimate | Std. Error | z value | Pr(>\|z\|) | |
| (Intercept) | 37.638361 | 5.876399 | 6.405 | 0.0000000001504 | *** | (Intercept) | 32.554388 | 6.144945 | 5.298 | 1.17E-07 | *** |
| MIN_Speed_Diff | -0.095534 | 0.015099 | -6.327 | 0.0000000002498 | *** | MIN_Speed_Diff | -0.05591 | 0.020027 | -2.792 | 0.005242 | ** |
| MAX_Speed_Diff | 2.861064 | 0.445999 | 6.415 | 0.0000000001409 | *** | MAX_Speed_Diff | 2.432157 | 0.470467 | 5.17 | 2.35E-07 | *** |
| MAX_Event_Speed | 0.026445 | 0.001967 | 13.446 | < 2e-16 | *** | MAX_Event_Speed | 0.024633 | 0.002078 | 11.855 | < 2e-16 | *** |
| STD_Event_Speed | -0.103621 | 0.010318 | -10.043 | <2e-16 | *** | STD_Event_Speed | -0.089453 | 0.011272 | -7.936 | 2.10E-15 | *** |
| Right_Exits | 0.263377 | 0.061958 | 4.251 | 0.0000212902784 | *** | Right_Exits | 0.225168 | 0.063591 | 3.541 | 0.000399 | *** |
| Outgoing_Lanes | -0.101075 | 0.04165 | -2.427 | 0.015200 | * | Outgoing_Lanes | -0.042743 | 0.046657 | -0.916 | 0.359619 | |
| Sideway | 0.493532 | 0.074886 | 6.59 | 0.0000000000439 | *** | Sideway | 0.568168 | 0.078439 | 7.243 | 4.37E-13 | *** |
| --- | | | | | | | | | | | |
| Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 | | | | | | | | | | | |

*Figure 4: Generalized linear models results*

The Poisson regression model identifies significant predictors of crash frequency at junctions using harsh acceleration events. The highly significant intercept term (3.901915) indicates a strong baseline crash frequency. Key speed-related variables include MIN_Speed_Diff (-0.104340) and MAX_Speed_Diff (0.082100), both statistically significant, with MIN_Speed_Diff reducing and MAX_Speed_Diff increasing crash frequency. MAX_Event_Speed (0.018499) is positively associated with crash frequency, while STD_Event_Speed (-0.128158) shows a negative association. Junction characteristics like Right_Exits (0.299347) are linked to higher crash frequency, whereas more Outgoing_Lanes (-0.126292) correlate with fewer crashes. Sideway (0.046146) is not statistically significant (p = 0.182).

Model diagnostics show a null deviance of 1435.25 and a residual deviance of 461.27, indicating the model explains significant variability in crash frequency. An AIC value of 696.21 suggests a good model fit, with a dispersion parameter of 1, indicating no overdispersion. Thus, the Poisson regression model highlights key predictors of crash frequency, informing strategies to reduce crashes by targeting specific junction characteristics and traffic dynamics.

The GLMER with random intercepts provides additional insights. The highly significant intercept estimate (3.609952) indicates a strong baseline crash frequency. MIN_Speed_Diff has a negative estimate (-0.082331), suggesting fewer crashes with increased minimum speed difference. Conversely, MAX_Speed_Diff has a positive estimate (0.076991), indicating more crashes with higher maximum speed differences. MAX_Event_Speed shows a positive estimate (0.015155), linking higher speeds during events to increased crash frequency. STD_Event_Speed has a significant negative estimate (-0.108751), implying fewer crashes with greater speed variability. Right_Exits has a positive estimate (0.198172), indicating higher crash frequencies at junctions with more right exits. Outgoing_Lanes has a non-significant estimate (0.003163), suggesting no significant impact on crash frequency. Sideway has a significant positive estimate (0.119132), indicating a positive association with crash frequency.

In conclusion, the GLMER model with random intercepts identifies significant predictors of crash frequency at junctions, including speed differences, maximum event speed, and standard deviation of event speeds. These findings highlight crucial factors for traffic management strategies aimed at reducing crashes.

Following the analysis of harsh acceleration events, the same methodology was used to predict the frequency of harsh braking events using various predictor variables. The intercept is estimated at 37.638361 with a standard error of 5.876399, indicating a strong baseline frequency of harsh braking events when all predictors are zero.

Among the key predictor variables, MIN_Speed_Diff has an estimate of -0.095534 with a standard error of 0.015099, which is highly significant, suggesting that an increase in the minimum speed difference is associated with a decrease in harsh braking events. Conversely, MAX_Speed_Diff has a positive estimate of 2.861064 with a standard error of 0.445999, indicating that higher maximum speed differences lead to an increase in harsh braking events. MAX_Event_Speed has an estimate of 0.026445 with a standard error of 0.001967, highly significant, showing that higher maximum speeds during events are linked to more frequent harsh braking events. In contrast, STD_Event_Speed has a negative estimate of -0.103621 with a standard error of 0.010318, also highly significant, indicating that greater speed variability is associated with fewer harsh braking events.

The presence of right exits is associated with an increase in harsh braking events, with an estimate of 0.263377 and a standard error of 0.061958, highly significant. On the other hand, the number of outgoing lanes has a negative estimate of -0.101075 with a standard error of 0.04165, which is significant, suggesting that more outgoing lanes are associated with fewer harsh braking events. Additionally, the presence of a sideway has a significant positive impact on the frequency of harsh braking events, with an estimate of 0.493532 and a standard error of 0.074886.

In conclusion, the GLM identifies several key predictors of harsh braking events, including speed-related factors and junction characteristics. The GLMER with random intercepts provides further insights into predictors of harsh braking events, accounting for variability across different junction types. The intercept estimate is 32.554388 with a standard error of 6.144945, indicating a substantial baseline frequency of harsh braking events when all other predictors are zero. Both models identify critical predictors of harsh braking events, providing valuable insights for traffic management strategies aimed at reducing such events by focusing on speed differences, maximum event speed, and junction characteristics.

For MIN_Speed_Diff, the estimate is -0.05591 with a standard error of 0.020027, suggesting that increased minimum speed difference is associated with fewer harsh braking events. MAX_Speed_Diff has a positive estimate of 2.432157 with a standard error of 0.470467, indicating that higher maximum speed differences lead to more harsh braking events.

MAX_Event_Speed has an estimate of 0.024633 with a standard error of 0.002078, demonstrating that higher speeds during events are linked to more frequent harsh braking events. STD_Event_Speed has a negative estimate of -0.089453 with a standard error of 0.011272, highly significant, suggesting that greater speed variability is associated with fewer harsh braking events.

Right_Exits has a positive estimate of 0.225168 with a standard error of 0.063591, indicating that more right exits are associated with an increase in harsh braking events. Conversely, Outgoing_Lanes has a negative estimate of -0.042743 with a standard error of 0.046657, suggesting no strong evidence of its impact on harsh braking events. Sideway has a significant positive impact, with an estimate of 0.568168 and a standard error of 0.078439.

In conclusion, the GLMER model with random intercepts highlights several significant predictors of harsh braking events, including speed differences, maximum event speed, variability in event speeds, and junction characteristics like the presence of right exits and sideway. These findings underscore the importance of considering these factors in traffic management and safety strategies to mitigate the frequency of harsh braking events.

| **Exploring the relationship between unsafe traffic events and crash occurrence using harsh acceleration events** | | | | | | |
|---|---|---|---|---|---|---|
| VIF scores for the Base GLM model: | | | | | | |
| MIN_Speed_Diff | MAX_Speed_Diff | MAX_Event_Speed | STD_Event_Speed | Right_Exits | Outgoing_Lanes | Sideway |
| 2.036474 | 1.853465 | 1.772092 | 1.77632 | 1.768705 | 3.948076 | 1.465606 |
| VIF scores for the Random intercepts GLMER model: | | | | | | |
| MIN_Speed_Diff | MAX_Speed_Diff | MAX_Event_Speed | STD_Event_Speed | Right_Exits | Outgoing_Lanes | Sideway |
| 1.903179 | 1.860567 | 1.868247 | 1.843664 | 2.180624 | 3.54389 | 1.219679 |
| **Exploring the relationship between unsafe traffic events and crash occurrence using harsh braking events** | | | | | | |
| VIF scores for the Base GLM model: | | | | | | |
| MIN_Speed_Diff | MAX_Speed_Diff | MAX_Event_Speed | STD_Event_Speed | Right_Exits | Outgoing_Lanes | Sideway |
| 2.162926 | 1.106514 | 1.813739 | 1.725633 | 2.031507 | 2.88136 | 1.404447 |
| VIF scores for the Random intercepts GLMER model: | | | | | | |
| MIN_Speed_Diff | MAX_Speed_Diff | MAX_Event_Speed | STD_Event_Speed | Right_Exits | Outgoing_Lanes | Sideway |
| 1.173837 | 1.076654 | 1.991475 | 2.006422 | 2.036128 | 1.96741 | 1.451817 |

*Figure 5: VIF scores results.*

The VIF scores provide insights into multicollinearity among predictors in both the base GLM and the GLMER with random intercepts. For acceleration events (see **Error! Reference source not found.**, upper side), the base GLM model shows MIN_Speed_Diff with a VIF of 2.036474, indicating moderate multicollinearity. Other variables, including MAX_Speed_Diff (1.853465), MAX_Event_Speed (1.772092), STD_Event_Speed (1.77632), Right_Exits (1.768705), and Sideway (1.465606), exhibit low multicollinearity. Outgoing_Lanes has a higher VIF of 3.948076, indicating moderate to high multicollinearity. The

GLMER model generally shows lower VIF scores: MIN_Speed_Diff (1.903179), MAX_Speed_Diff (1.860567), MAX_Event_Speed (1.868247), and STD_Event_Speed (1.843664) indicate low multicollinearity, while Right_Exits (2.180624) and Outgoing_Lanes (3.54389) suggest moderate to high multicollinearity. Sideway has the lowest VIF at 1.219679, indicating very low multicollinearity.

For harsh braking events (see **Error! Reference source not found.**, bottom side), the base GLM model shows MIN_Speed_Diff with a VIF of 2.162926, indicating moderate multicollinearity, and MAX_Speed_Diff with a VIF of 1.106514, suggesting very low multicollinearity. MAX_Event_Speed (1.813739) and STD_Event_Speed (1.725633) show low multicollinearity, while Right_Exits (2.031507) and Outgoing_Lanes (2.88136) indicate moderate multicollinearity. Sideway has the lowest VIF at 1.404447, indicating low multicollinearity. In the GLMER model, VIF scores are generally lower: MIN_Speed_Diff (1.173837), MAX_Speed_Diff (1.076654), MAX_Event_Speed (1.991475), STD_Event_Speed (2.006422), and Outgoing_Lanes (1.96741) indicate low multicollinearity. Right_Exits (2.036128) shows moderate multicollinearity, while Sideway (1.451817) has low multicollinearity.

In summary, the VIF scores indicate low to moderate multicollinearity for most predictors in both models, with Outgoing_Lanes showing the highest multicollinearity in harsh braking events, and slightly lower VIF scores in the GLMER model suggesting that random intercepts help mitigate multicollinearity, thereby affirming the robustness of the models and the reliability of the predictors in explaining crash frequency at junctions for harsh accelerating and braking events.

| | | | | | | |
|---|---|---|---|---|---|---|
| **Exploring the relationship between unsafe traffic events and crash occurrence using harsh acceleration events** | | | | | | |
| LRT results: | | | | | | |
| Likelihood ratio test | | | | | | |
| Model 1: FREQUENCY_ACC ~ MIN_Speed_Diff + MAX_Speed_Diff + MAX_Event_Speed + STD_Event_Speed + Right_Exits + Outgoing_Lanes + Sideway | | | | | | |
| Model 2: FREQUENCY_ACC ~ MIN_Speed_Diff + MAX_Speed_Diff + MAX_Event_Speed + STD_Event_Speed + Right_Exits + Outgoing_Lanes + Sideway + (1 \| Junct_Type) | | | | | | |
| | #Df | LogLik | Df | Chisq | Pr(>Chisq) | |
| 1 | 8 | -340.1 | | | | |
| 2 | 9 | -318.01 | 1 | 44.185 | 0.00000000003 | *** |
| Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 | | | | | | |
| LRT = significant, so GLMER is better | | | | | | |
| **Exploring the relationship between unsafe traffic events and crash occurrence using harsh braking events** | | | | | | |
| Likelihood ratio test | | | | | | |
| Model 1: FREQUENCY_BRK ~ MIN_Speed_Diff + MAX_Speed_Diff + MAX_Event_Speed + STD_Event_Speed + Right_Exits + Outgoing_Lanes + Sideway | | | | | | |
| Model 2: FREQUENCY_BRK ~ MIN_Speed_Diff + MAX_Speed_Diff + MAX_Event_Speed + STD_Event_Speed + Right_Exits + Outgoing_Lanes + Sideway + (1 \| Junct_Type) | | | | | | |
| | #Df | LogLik | Df | Chisq | Pr(>Chisq) | |
| 1 | 8 | -138.31 | | | | |
| 2 | 9 | -135.82 | 1 | 4.9793 | 0.02565 | * |
| Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 | | | | | | |
| LRT = significant, so GLMER is better | | | | | | |

*Figure 6: Likelihood Ratio Test (LRT) results.*

The results of the LRT provide a comparison between the base GLM (Model 1) and the GLMM (Model 2) with random intercepts for junction type, conducting both for harsh accelerating(upper side) and braking events(bottom side), as shown in **Error! Reference source not found.**.

<u>For harsh accelerating events:</u> In Model 1, the predictors include MIN_Speed_Diff, MAX_Speed_Diff, MAX_Event_Speed, STD_Event_Speed, Right_Exits, Outgoing_Lanes, and Sideway. This model yields a log-likelihood of -340.1. Model 2 includes the same predictors as Model 1, but also incorporates a random intercept for junction type. This model achieves a log-likelihood of -318.01. The difference in degrees of freedom (Df) between the two models is 1. The chi-squared value for the test is 44.185, with a corresponding p-value of 0.00000000003, which is highly significant ($p < 0.001$).

The significant result of the LRT indicates that the inclusion of a random intercept for junction type significantly improves the model fit. Therefore, the GLMER model (Model 2) is statistically superior to the base GLM model (Model 1). This finding suggests that accounting for variability in accident frequency across different junction types provides a better understanding of the factors influencing accident occurrences.

<u>For harsh braking events:</u> Model 1, which includes the predictors MIN_Speed_Diff, MAX_Speed_Diff, MAX_Event_Speed, STD_Event_Speed, Right_Exits, Outgoing_Lanes, and Sideway, yields a log-likelihood of -138.31. Model 2 includes the same predictors as Model 1 but adds a random intercept for junction type (Junct_Type), resulting in a log-likelihood of -135.82. The difference in Df between the two models is 1. The chi-squared value for the test is 4.9793, with a corresponding p-value of 0.02565, which is significant at the 5% level.

The significant result of the LRT indicates that including a random intercept for junction type significantly improves the model fit. Therefore, the GLMER model (Model

2) is statistically superior to the base GLM model (Model 1). This finding suggests that accounting for variability in harsh braking frequency across different junction types provides a better understanding of the factors influencing these events, highlighting the importance of incorporating random effects in the model.

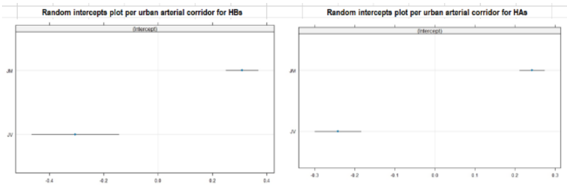


*Figure 7:Random intercepts plot per urban arterial corridor.*

The **Error! Reference source not found.** displays random intercepts plots for urban arterial corridors (JM and JV) in two models: harsh braking events (HBs) on the left and harsh acceleration events (HAs) on the right. For HBs, corridor JM has an intercept slightly above zero, indicating a higher baseline frequency of harsh braking events compared to the overall average, while corridor JV has an intercept below zero, suggesting a lower baseline frequency. Similarly, for HAs, corridor JM shows a higher baseline frequency (intercept above zero) and corridor JV a lower one (intercept below zero). This variability underscores the importance of including random effects in the models to capture unique characteristics of each corridor, providing more precise insights into factors influencing harsh braking and acceleration events and enhancing the understanding of traffic dynamics and safety measures.

## 4. Conclusion, Limitations and Future Research

This study assessed the impact of harsh acceleration and braking events on crash frequency at junctions on two urban expressways in Athens. Using high-resolution smartphone data from 303 drivers, traffic data from 26 loops, and Google Maps Road characteristics, the study evaluated GLM and GLMM.

The GLM showed significant predictors with an AIC of 292.63 and residual deviance of 106.87 on 25 degrees of freedom. Adding a random intercept for junction type in the GLMM improved model fit, with a lower AIC (289.6) and better log-likelihood (-135.82). The log-likelihood ratio test confirmed the GLMM's superior fit (p = 0.02565). Significant predictorsincludedMIN_Speed_Diff,MAX_Speed_Diff, MAX_Event_Speed, STD_Event_Speed, Right_Exits, and Sideway. The caterpillar plot indicated higher baseline crash frequencies for junction type JM compared to JV. VIF values below 5 indicated low multicollinearity.

Findings suggest practical safety measures: reducing speed variability through consistent limits and enforcement, enhancing driver awareness, and improving junction design with clearer signage, dedicated right-turn lanes, and better sight lines. Well-marked sideway areas for pedestrians and cyclists can also enhance safety.

The study highlights the importance of considering both fixed and random effects in traffic safety analysis. The higher crash risk at junction type JM suggests interventions like targeted enforcement and junction redesign. Future research should collect more detailed data on driver behavior, traffic flow, and environmental conditions, and use advanced modeling techniques like machine learning to gain deeper insights. Addressing limitations such as the small sample size of 33 junctions, data accuracy issues, and excluded factors like weather, demographics, and vehicle types can enhance applicability across different contexts.

## 5. References

Abdel-Aty, M., Uddin, N., Pande, A., Abdalla, F. M., & Hsia, L. (2013). Predicting crashes on expressway ramps using statistical methods, rough sets, and neural networks. Journal of Transportation Engineering, 139(5), 481-492.

Archer, J. (2005). Indicators for Traffic Safety Assessment and Prediction and Their Application in Micro-Simulation Modelling: A Study of Urban and Suburban Intersections. Royal Institute of Technology.

Barić, D., Živković, M., & Štimac, I. (2014). The Influence of the Use of Smartphones on Driver Behaviour and Traffic Safety. Transport Research Arena.

Bonsall, P. W., Liu, R., & Young, W. (2005). Modelling safety-related driving behaviour—impact of parameter values. Transportation Research Part A: Policy and Practice, 39(5), 425-444.

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. Proceedings of the National Academy of Sciences, 113(10), 2636-2641.
ELSTAT. (2020). Road traffic crashes. Available at Hellenic Statistical Authority (ELSTAT), 2020.
European Commission (2018). Annual Accident Report 2018. Directorate General for Mobility and Transport.
European Commission (2019). Road Safety in the European Union: Trends, Statistics and Main Challenges.
European Commission. (2023). Road safety in the EU: fatalities below pre-pandemic levels but progress remains too slow. Available at: European Commission Road Safety Report 2023
European Commission. (2024). Latest key figures. Available at: European Road Safety Observatory
Eurostat. (2023). Road fatalities up 4% in 2022. Available at: Eurostat Road Fatalities 2022 (European Commission), Eurostat Road Fatalities 2022
Hamdar, S. H., Mahmassani, H. S., & Treiber, M. (2016). From behavioral psychology to acceleration modeling: Calibration, validation, and exploration of drivers' cognitive and safety parameters in a risk-taking environment. Transportation Research Part B: Methodological, 78, 32-53.
Hanowski, R. J., Olson, R. L., Hickman, J. S., & Dingus, T. A. (2005). The 100-car naturalistic driving study: A descriptive analysis of light vehicle-heavy vehicle interactions from the light vehicle driver's perspective.
Hauer, E. (1997). Observational Before-After Studies in Road Safety. Pergamon Press.
Hellenic Statistical Authority (ELSTAT), 2020.
Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. Accident Analysis & Prevention, 38(1), 185-191.
Jansen, J. J., & Wesseling, A. (2018). Behavioral patterns of young drivers during sudden, simulated crashes: A driving simulator study. Traffic Injury Prevention, 19(2), 156-161.
Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2009). The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. National Highway Traffic Safety Administration.
Laureshyn, A. (2010). Application of Automated Video Analysis for Behavioral Studies: Conceptual Framework and Practical Use in Traffic Safety. Lund University.
Mantouka, E., Vlahogianni, E. I., & Golias, J. C. (2018). Real-time detection of dangerous events in urban arterials using high-resolution traffic data. Transportation Research Part C: Emerging Technologies, 86, 270-278.
Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). Driver distraction in commercial vehicle operations. National Highway Traffic Safety Administration, Final Report DOT HS 811 043.
Paefgen, J., Staake, T., & Fleisch, E. (2014). Multivariate exposure modeling of accident risk: Insights from pay-as-you-drive insurance data. Transportation Research Part A: Policy and Practice, 61, 27-40.
Papadimitriou, E., Yannis, G., & Folla, K. (2018). A review of driving performance parameters investigated in the context of naturalistic driving studies. International Journal of Injury Control and Safety Promotion, 25(1), 9-19.
Roetting, M., Huang, Y. H., McDevitt, J. R., & Melton, D. (2003). When technology tells you how you drive–truck drivers' attitudes towards feedback by technology. Transportation Research Part F: Traffic Psychology and Behaviour, 6(4), 275-287.
Petraki V., Ziakopoulos A., Yannis G. (2020) Combined impact of road and traffic characteristic on driver behavior using smartphone sensor data, Accident Analysis and Prevention 144 (2020) 105657.
Shaout, A., & Bodenmiller, R. (2011). Evaluation of safe driving behavior using smartphone sensors. International Journal of Advanced Computer Science and Applications, 2(8), 19-24.
Singh, S. (2015). Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey. National Highway Traffic Safety Administration, Report No. DOT HS 812 115.
Singh, S. (2018). Driver behavior and crash risk: Identifying and mitigating driver errors. Journal of Safety Research, 66, 165-173.
Stephens, A. N., & Groeger, J. A. (2009). Anger-congruent behaviour in simulated driving. Transportation Research Part F: Traffic Psychology and Behaviour, 12(5), 382-388.
Toledo, T., Musicant, O., & Lotan, T. (2008). In-vehicle data recorders for monitoring and feedback on drivers' behavior. Transportation Research Part C: Emerging Technologies, 16(3), 320-331.
Troncoso, M. C., Efthymiou, A. K., Beligiannis, G. N., & Tzovaras, D. (2010). Driver's risk index based on a novel fuzzy logic system for usage-based insurance. 2010 IEEE Intelligent Vehicles Symposium, 965-970.
Tselentis, D. I., Yannis, G., & Vlahogianni, E. I. (2017). Innovative insurance schemes: pay as/how you drive. Transportation Research Procedia, 25, 2104-2113.
Vlahogianni, E. I., & Barmpounakis, E. N. (2017). Driving analytics using smartphone data: Challenges and opportunities. Transportation Research Part C: Emerging Technologies, 79, 196-206.
WHO. (2018). Global status report on road safety 2018. Available at: WHO Global Status Report on Road Safety, WHO Report 2018
Wouters, P., & Bos, J. M. J. (2000). Traffic accident reduction by monitoring driver behaviour with in-car data recorders. Accident Analysis & Prevention, 32(5), 643-650.
Yamakado, T., Shimazaki, Y., & Miyajima, C. (2009). Eco-driving support system for manual transmission vehicles based on driving diagnosis. SAE Technical Paper No. 2009-01-1318.
Yannis, G., Papadimitriou, E., & Folla, K. (2014a). Effect of GDP changes on road traffic fatalities. Safety Science, 63, 42-49.
Yannis, G., Papadimitriou, E., & Folla, K. (2014b). Effect of GDP changes on road traffic fatalities. Journal of Safety Research, 50, 77-85.
Ziakopoulos, A., & Yannis, G. (2020a). A review of spatial approaches in road safety. Accident Analysis & Prevention, 135, 105327.
Ziakopoulos, A., & Yannis, G. (2020b). Meta-analysis of the effects of road safety measures on crashes by road type. Accident Analysis & Prevention, 141, 105508.
Zohar, D., Lee, J., & Drach-Zahavy, A. (2014). "Social comparison and driver behavior: The effect of random feedback in a simulated driving task." Transportation Research Part F: Traffic Psychology and Behaviour, 24, 67-77.
Zhao, Y., Zheng, J., & Zhang, Y. (2017). "Using Smartphone Sensors to Evaluate Safety Effects of Transit Signal Priority at Signalized Intersections." Journal of Intelligent Transportation Systems, 21(3), 232-242.