Accident Analysis and Prevention 144 (2020) 105657

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Combined impact of road and traffic characteristic on driver behavior using smartphone sensor data

Virginia Petraki, Apostolos Ziakopoulos*, George Yannis

National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Iroon Polytechniou St., GR-15773, Athens, Greece

ARTICLE INFO

Keywords:
Driver behavior
Harsh events
Geometric characteristics
Traffic characteristics
Smartphone data

ABSTRACT

The objective of this research is to exploit high resolution driving behavior data collected via sensors of smartphones from 303 drivers in order to examine driver behavior at road segment and junction level. These sensor data are combined with traffic and road geometry characteristics and subsequently depicted spatially using Geographical Information System software. Events of harsh driver behavior (8592 harsh accelerations and 3946 harsh brakings) were mapped to delimited segments and junctions of two urban expressways in Athens, Greece. For the analysis, two multiple linear regression models and two log-linear regression models were developed. Results indicate that in road segments there is an increase in the number of harsh events if average traffic flow per lane increases in the respective areas. Furthermore, as the average occupancy increases in junctions, there is an increase in harsh accelerations, and as the average speed increases, more harsh deceleration events occur. It is evident that traffic characteristics (traffic flow & speed) have the most statistically significant impact on the frequency of harsh events compared to factors related to road geometry and driver behavior.

1. Introduction

1.1. Road safety state

Despite the fact that several consistent attempts have been made to reduce road crash numbers and the respective casualties, it appears that fatalities from crashes have platooned during the recent years with the global number of deaths rising to 1.35 million annually (WHO, 2018) corresponding to around 3700 daily fatalities globally. Regarding Europe, according to official statistics, in 2016 25,600 fatal crashes were reported on the roads of the European Union amounting to about 70 deaths per day. More than 1.4 million people were injured from road crashes, amounting to about 3600 injuries per day (European Commission, 2018).

Greece, where the analyses of the present research take place, is the country with the highest crash fatality reduction from road crashes in Europe (51 %) between 2009 and 2018. However, with 64 deaths per million it still ranks 22nd among the 28 states of the European Union (European Commission, 2019). Furthermore, some of the observed crash reduction is expected to be part of the economic recession as annual GDP decreases have been shown as associated with mortality rate decreases (Yannis et al., 2014a). According to the Hellenic Statistical Authority (ELSTAT, 2020) in January 2018, road crashes that occurred throughout the country and caused the death or injury of people increased by 18.8 % compared to the equivalent of 2017 (757 in January 2018 compared to 637 in January 2017).

1.2. State of the art

1.2.1. The importance of driving behavior analysis

The analysis of driver behavior has been established as a critical part of preventing road crashes and improving road safety. While it is established that the three main factors of a road crash are human factors (driver/road user behavior), road environment/design faults and vehicle faults; driver behavior has been determined as the critical reason for about 95 % of total road crashes (Singh, 2015). The critical reason of a crash is defined by the NHTSA 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 comprises in turn a large number of factors that have been found to contribute to road crashes (Dingus et al., 2016). The road environment comprises of several different elements which may in turn influence driving behavior differently, as past research indicates (Horberry et al., 2006; Hamdar et al., 2016).

With the present state of road safety in mind, it is logical to seek alternative venues for crash reduction. The exploitation of new

* Corresponding author.
E-mail address: apziak@central.ntua.gr (A. Ziakopoulos).

https://doi.org/10.1016/j.aap.2020.105657
Received 4 November 2019; Received in revised form 12 June 2020; Accepted 16 June 2020
Available online 04 July 2020
0001-4575/

technological advancements allows for driver monitoring through smartphone applications and the respective data collection and processing (Vlahogianni and Barmpounakis, 2017). Smartphones have the advantage of being programmable and a wide array of sensors has now become standard equipment that can be utilized for transport studies (such as accelerometer, digital compass, gyroscope, GPS, microphone and camera) and enable sensing applications, even without user engagement (Mantouka et al., 2018).

Furthermore, the analysis of driving behavior seems to be useful in the car insurance market, while insurance is a significant proportion of car costs. To reduce costs for the car owner and for insurers, insurance companies have developed various programs such as Pay-As-You-Drive (PAYD). For instance, drivers-customers who use the PAYD program will be charged depending on the location and the time they drive against a specified amount of money that they would pay each year with another program (paid-by-the-year). In other words, the results of such an analysis will help to create new road safety programs or to reinforce existing ones. This way, the insurance programs will encourage responsible driving, reduce the risk of road crashes and save lives and money (Troncoso et al., 2010).

Aggressive driving behavior parameters, such as harsh accelerations and decelerations, and their correlations with crash risk, have been investigated by the insurance industry (Paefgen et al., 2014). Harsh events have been determined as strongly correlated with driving risk (Tselentis et al., 2017). Earlier studies have documented harsh driving behavior as critical for driving risk assessment (Bonsall et al., 2005).

Furthermore, harsh events have a significant impact on energy efficiency as well. The difference, under terms of fuel consumption and of gas emissions, between a safe (or calm) driver and an aggressive one is estimated to be greater than 40 % (Alessandrini et al., 2012). For this reason, with the intention of reducing the environmental footprint due to the road transport system, in recent years the idea of educating drivers to adopt a more environmentally friendly way of driving has been promoted. Such behavior could be achieved by reducing harsh accelerations, harsh decelerations and harsh maneuvers (Yamakado et al., 2009). The desired reduction in the emission of gases could be accomplished by finding the factors influencing aggressive behavior through its analysis.

1.2.2. Exploitation of sensor data

While analyzing driver behavior is crucial, the difficulty in collecting reliable and high-resolution data restricts progress in this field. The required data can be collected in different ways, such as (i) questionnaires that investigate a driver's personality and driving styles to understand his risk of being involved in a road crash, (ii) simulators where a driver controls a car in a virtual driving environment, providing a safe and isolated way to study driver behavior, (iii) in-vehicle data recorders that collect driving variables in real-time and in a naturalistic driving environment (e.g. OBD).

Some of the previous methods of collecting driving behavior data require high costs to be applied and may not yield objective results. The proliferation of smartphones and the various types of integrated sensors in them created a fourth method of collection: a cheap and easy to install platform for detecting driver behavior in naturalistic conditions, offering a low cost alternative to driving data collection. The evaluation of driver behavior through experiments using smartphone data was found to be a very promising method, enabling the acquisition of a wealth of real-life data on driving behavior and related risks such as distraction and speeding (Papadimitriou et al., 2018). The dissemination of modern technologies for mobile devices, as well as the development of several applications for the utilization of their internal sensors, allow users to interact by getting a feedback in real-time about their driving behavior. Something like that can be useful in reinforcing driver awareness and promoting safety. By providing such interventions during driving, a reduction of approximately 20 % is calculated in the average estimated number of road crashes under specific conditions (Wouters and Bos, 2000).

Due to the rapid technological progress, especially in telematics and Big Data analytics, along with increases in the information technologies' penetration and use by drivers (e.g. smartphones), studies examine the driving behavior through the collection of data using technology devices adapted to the brain of the vehicle or through sensors of a smartphone.

There are several measurements which can be used alone or in combination in order to evaluate the driving behavior. These may be related to the way of driving, the mechanical characteristics of the vehicle, the weather conditions, the duration of the journey, the distance travelled, etc. In scientific literature as the most significant parameters for unsafe or aggressive driving style assessment appears to be longitudinal and lateral accelerations and decelerations (Klauer et al., 2009; Shaout and Bodenmiller, 2011; Paefgen et al., 2012; Vaiana et al., 2014). More specifically, significant different frequencies of events were observed depending on the type of road in which event was occurred. That observation can be explained due to the traffic in the city which requires more accelerations, decelerations and sharper turns (Paefgen et al., 2012). Furthermore, in another investigation there were detected more aggressive behaviors in the area that was characterized by high traffic flow and located in the city center, while fewer were observed in the area that could be classified as suburban (Vaiana et al., 2014).

The recording of driving behavior in naturalistic conditions, either from smartphone sensors or from in-vehicle device sensors, presents advantages such as the collection of high-resolution data and objective measurements, as well as the possibility of informing drivers about their driving style by receiving feedbacks. A noteworthy fact is that exposure to feedback is an important incentive to drive more carefully and in a more responsible manner. Research has shown that drivers accept more easily feedback derived from technological methods compared to other methods (Roetting et al., 2003; Toledo et al., 2008).

Significant correlations are found between the vehicle-related recording systems and the sensors of smartphones, although they are affected from the type of event, the location of the smartphone in the car and from external factors (Paefgen et al., 2012). However, due to the high cost of installing and operating a system unit in the vehicle for data collection, the creation of an application for smartphones which provides data by exploiting their sensors is more strongly supported.

Several studies have been carried out focusing on the spatial analysis of recorded road crashes, road characteristics and census variables, and ultimately on the export of road crash models to improve road safety. Due to the progress of the geographic information systems it is possible the analysis of road crashes can be conducted across different geographic units. However, there is no clear guideline on which geographic unit should be selected for the spatial analysis of road crashes (Ziakopoulos and Yannis, 2020a). Differentiations across larger spatial units have been shown to influence coefficient results via meta-regression techniques (Ziakopoulos and Yannis, 2020b). The preference of the spatial unit may vary depending on the dependent variable of the mathematical model, and furthermore, it has been found that the distance travelled by the vehicle and the number of junctions in the area under study constitute statistically significant variables for the analysis of road crashes (Abdel-Aty et al., 2013).

1.3. Merits of harsh event analysis

Since road crashes or road crash casualties are a traditional focus in the science of road safety, one might argue: What is the value of analyzing harsh events?

Harsh events have been adopted as a parameter for measurement of road safety in the past, as they are strongly correlated with reduced spatial and temporal headways (unsafe distance) from neighboring vehicles, near misses with road users or stationary objects, and also include additional behavioral parameters such as lack of concentration or experience. Harsh events have been determined as inherently linked with driving risk (Tselentis et al., 2017), while research has also documented harsh driving behavior as critical for driving risk assessment (Bonsall et al., 2005; Gündüz et al., 2017). Harsh accelerations and decelerations, and their correlations with crash risk, have been investigated by the insurance industry as well (Paefgen et al., 2014).

While harsh events are ultimately driver behavior metrics, they have the potential to be analyzed as point-data (locations), much like road crashes. Consequently, the examination of patterns in the distribution of harsh event points does have the potential to reveal interesting underlying mathematical relationships that show dependencies with independent parameters. An aggressive driver will have elevated harsh events not only in a particular trip, but in all trips made across the map. Thus a large enough driver sample, leading to a sizeable trip sample, can be reasonably expected to convey useful information about problematic road segments with high road safety risk (hotspots). Moreover, harsh events constitute pro-active road safety parameters and can thus disclose these hotspots preemptively, before crashes occur and their respective consequences manifest.

Furthermore, harsh events are expected to be increasingly employed as an important driver classification metric in usage-based motor insurance (UBI), as they appear to be more representative of crash occurrence probability (Tselentis et al., 2017). However, to the experience of the authors, studies focusing on factors influencing harsh event occurrence and similar characteristics are significantly outnumbered by studies analyzing crashes, indicating significant research gaps in this field.

At this point it should be underlined that harsh accelerations and harsh brakings are two different phenomena occurring during different situations. As such, it is recommended that they are not analyzed collectively in principle. Indicatively, drivers with higher anger, frustration and anxiety levels display higher acceleration values and apply increased physical pressure on the accelerator pedal (Stephens and Groeger, 2009). Harsh braking events are thought to indicate reaction in anticipation of a safety-critical event (e.g. near-miss) or crash, and are used as indicators for that purpose in naturalistic driving data (e.g., Hanowski et al., 2005; Olson et al., 2009; Zohar et al., 2014; Jansen and Wesseling, 2018).

1.4. Aim of the current research

In light of the aforementioned aims and methods, the aim of the current research is the utilization of high resolution smartphone data for the examination of frequencies of harsh accelerations and harsh decelerations (brakings) and their concentrations across junctions and road segments. The combined impact of road and traffic characteristics on driver behavior using data from smartphones will be investigated using multi-source data. To the extent of the authors' knowledge, this is an unexplored research direction that can offer valuable insights to understanding the factors affecting driver behavior as described by harsh events on a mesoscopic/macroscopic road segment basis.

2. Data collection and processing

2.1. Data collection

The study area consisted of two urban expressways that were examined, Mesogeion Avenue and Vouliagmenis Avenue. Both avenues are in Athens and where selected mainly due to the comparable number of traffic lanes and the separation of the two directions that they feature. The two Avenues appear on Fig. 1 .

To enable the analysis of the aggressive driving behavior, data from three separate sources were analyzed. Specifically, driving behavior data, traffic characteristics of the two urban expressways under study as well as road geometry characteristics were collected. The first dataset concerned the driving behavior of 303 drivers in Athens and was collected using sensors of smartphones via the purpose-made application of OSeven telematics, which is a driver-friendly telematics application. The second one consisted of traffic characteristics which were collected through twenty-six (26) inductive loops, installed by the Traffic Management Centre of Attica Region in specific measuring positions on the two urban expressways under study. Finally, the third dataset was formed by geometry characteristics of the two urban expressways which were collected using the online mapping service provided by Google Maps. The data collection process is described in more detail in the following sections.

An innovative data collection scheme using a smartphone application that has been developed by OSeven Telematics was exploited for the recording and collection of driving behavior data. This is an integrated system for the individualized recording, collection, storage, evaluation and presentation of driving behavior data from smartphone sensors. Recorded data come from various smartphone sensors (accelerometer, gyroscope, magnetometer and GPS) and data fusion algorithms provided by Android and iOS. Advanced machine learning (ML) algorithms are used to remove outliers, clean and normalize the data. This ML output can be then exploited and used as basis for statistical analyses in independent research.

The application also detects harsh events which are analysed in this research, using data from all the axes of the accelerometer. The harsh events are calculated via data fusion and machine learning algorithms and not a rule based approach using as input the values of the accelerometer as well as values from additional sensors (e.g. orientation, GPS, gyroscope). Therefore, there is not a specific threshold of the accelerometer value for the determination of the harsh events, but rather accelerometer spike detection rules.

The smartphone application has been utilized for road safety research and described in past studies as well (Yannis et al., 2017; Tselentis et al., 2018, 2019; Stavrakaki et al., 2019). A comparable approach is followed in the present paper. A significant amount of data is recorded using this platform, as described in recent research that has utilized this specific scheme (Papadimitriou et al., 2019). The outputs of the OSeven algorithms have been evaluated both in the published studies and and used by major insurance companies in several countries (Brazil, Greece, Qatar, UK); this serves as evidence regarding the acceptance of the OSeven algorithms.

It is important to mention that data are completely anonymized before being provided by OSeven so that driving behavior of each application user cannot be connected with any personal information (thus driver characteristics such as age, gender etc. are not used in this study), in accordance with standing European personal data protection laws (GDPR). This application involves a data exploitation approach that is not user intrusive. Participant drivers had the application installed in their own smartphone devices; the application supports both Android & iOS devices. The OSeven algorithms and platform are device-agnostic. This means that all the algorithms have been calibrated in order to produce comparable results regardless of the device used for data collection. The smartphone devices include sensors of very high accuracy. Using advanced filtering algorithms, OSeven clears the noise so that the results are equivalent with the results that are produced between devices across manufacturers, operating systems (Android, iOS), as well as those results provided by other instrumentation approaches (e.g. OBD). No further information is provided on user mobile phones due to standing GDPR laws. The flexible smartphone installation and transferability of the application enables the acquisition of data for several parameters while enabling researchers to eschew vehicle instrumentation or video examination for eye-tracking movements which involve increased effort and costs (Ziakopoulos et al., 2020).

In total, 303 drivers participated in the smartphone naturalistic driving experiment in Athens and between 25 August 2016 to 26 November 2017 leading to the creation of two large databases of harsh accelerations and decelerations with thousands of events each. More specifically, during this period 4869 harsh accelerations and 2181

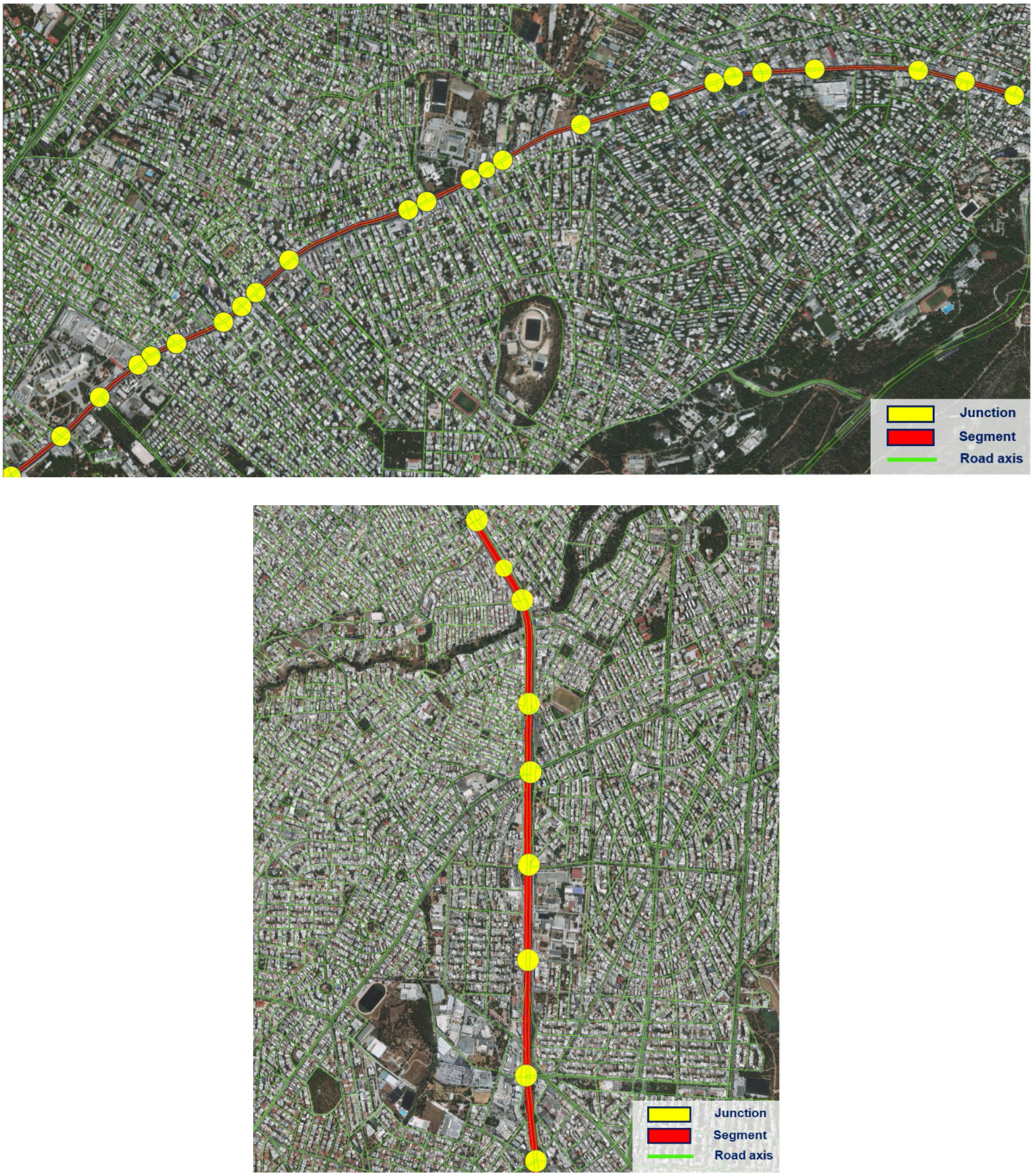

Fig. 1. Research area: Mesogion Ave. (above) & Vouliagmenis Ave. (below).

harsh brakings were recorded in Mesogeion Avenue and 3723 harsh accelerations and 1765 harsh brakings were recorded in Vouliagmenis Avenue. Event intensity is also classified among different categories. For each harsh event that was recorded, respective descriptive variables (speed of the event, maximum speed difference in two seconds during the event (a form of range), distance of the event) were recorded as well. The naturalistic approach adopted in this study entails no route or time limitations or other interventions imposed upon participant drivers. In other words, the analysis of harsh events was based exclusively on data from each passage from the study areas regardless of time. Additionally, events were analyzed without examining whether the passage from each segment was part of a greater trip or not.

Traffic characteristics were collected using the infrastructure provided by the Traffic Management Centre (TMC) of Attica Region. Specifically, 26 available measuring positions of traffic flow and occupancy were found in Vouliagmenis Ave. and Mesogeion Ave., installed by the TMC. However, as far as Vouliagmenis Ave. is concerned, only the measuring positions between Agiou Constantinou and Alimou Ave. taken into consideration due to the low quality of the remaining measurement locations. The obtained traffic data consists of a group of vehicles and are collected on two time bases. The first group is consisted of measurements at intervals of 90 s and, through their aggregation, the second group is obtained, which consists of hourly measurements. The hourly database referring to each typical and indicative day of the year 2017 was used for the purposes of the current research. TMC infrastructure and obtained data have also been utilized in past road safety research (Yannis et al., 2014b; Theofilatos, 2017; Theofilatos et al., 2017; Theofilatos et al., 2018).

It was decided that the utilized traffic parameters would be the average hourly traffic flow (normalized as average hourly traffic flow/lane, measured in vehicles/hour), average occupancy (measured as absolute percentage) and average aggregated speed (measured in kilometers/hour). The yearly average of 2017 was calculated at each measurement location for these parameters in order to acquire the typical value for each parameter (essentially resulting in AADT/lane for traffic flow, and yearly averages of hour-level values for occupancy and aggregated speed).

In order to be able to assess the influence of the road design on the aggressive behavior of the driver sample, the constructional configuration of each junction and road segment was determined through examination of aerial photography. Specifically, the geometrical characteristics of the two urban expressways were acquired with the help of the online mapping service Google Maps. The following variables were collected:

- number of entrances and exits of each junction under study
- number of outgoing and ingoing traffic lanes to and from the junction
- presence or absence of access roads
- number of right exits and entrances of the road segment
- presence or absence of bus lanes

2.2. Data processing

After driving behavior data, traffic characteristics and road geometry characteristics were obtained, they underwent considerable manipulation and processing. The first step consisted of trimming the harsh events from the entirety of the road network to the segments and junctions of the examined urban expressways. The next step consisted of the correct matching of traffic and geometry data to the corresponding segments and junctions that they characterized.

The analysis of the collected data was conducted on a segment basis with the geographical information software system (ArcGIS 10.3 GIS Software, 2020) and specifically the application ArcMap (ArcMap 10.3 GIS Software, 2020). The various stages of data processing via GIS are visualized on Fig. 2.

Initially, the road axes of Athens were added and after that the harsh accelerations and decelerations were imported as points, based on their coordinates. In addition, the map was imported to facilitate the easier and faster detection of the correct road segments. In order to classify the harsh events in junctions and road segments according to the spatial unit that they occurred into, the two urban expressways under study were detected on the map and afterwards circles of 50 m radius were designed in place of junctions, accompanied by 40 m radius cycles for pedestrian crossings and polygons for road segments on the expressways. Pedestrian crossings were too few in order to be meaningfully included in the analysis in the study areas.

On that note, the intersection safety literature lacks a commonly accepted threshold of intersection influence area. Abdel-Aty and Wang (2006)consider crashes within 250 ft (76.2 m) of an intersection milepost as 'at intersection'. Huang et al. (2013) have analyzed videos from cameras fixed at an intersection and installed 100 m before that intersection. More recently, Essa and Sayed (2018) and Guo et al. (2019) have used overlapping surfaces of varying sizes to monitor conflicts at an intersection via video analysis.

The choice of 50 m was selected as reasonable for the study areas at hand. This is because the two investigated Avenues feature a high density of small and medium-sized intersections and there was a need to avoid large continuous overlap which would be unrealistic for this research.

A challenge to be tackled at this point was the determination of the direction that each harsh event occurred on, for its assignment on the respective road segment. It should be noted that the smartphone sensor/GPS accuracy is a minimum of +/- 5 m, therefore there might be small inaccuracies in the mapping of events, especially regarding their direction (by being recorded in the opposite direction). To conduct the mapping, polygons that coincided with the road segments of each traffic direction with a width equal to the sum of the widths of the traffic lanes of the particular direction plus the half width of the median strip were drawn. It was assumed that the overall smartphone sensor/GPS inaccuracies would essentially cancel themselves out, leaving the resulting allocation of events per direction close to reality.

In this way, 17 road segments per direction and 24 junctions were defined on Mesogeion Ave., and 24 road segments per direction and 26 junctions were defined on Vouliagmenis Ave. As far as Vouliagmenis Ave. is concerned, not all road segments and junctions were taken into consideration due to the low quality of traffic measurements on certain parts of the expressway as previously described. The remaining junctions and the road segments were numbered during their design process so that the later data analysis would avoid any errors regarding the location of the spatial unit on the urban expressway.

In order to achieve the categorization of the harsh events in each numbered spatial unit that was defined, a geoprocessing model was used. In this way each spatial unit was now characterized by the frequency of harsh acceleration or harsh deceleration and by the variables that characterized any harsh event. Since the harsh events were not analyzed in isolation but as total depending on where they occurred, values of the descriptive variables (minimum number, maximum number, standard deviation, range, mean) of the harsh events that occurred in the same spatial unit were calculated through the geoprocessing model.

Finally, in order to combine the driving behavior high resolution data from the intelligent mobile phone sensors and the traffic sizes measured by the inductive loops placed by the TMC, the traffic measurement locations were added in the map under processing, based on their Cartesian coordinates. Consequently, a dataset was produced that comprised the events of harsh accelerations and decelerations in Athens, the geometric characteristics of the road network and the measurement positions of traffic characteristics, as shown on Table 1.

Due to the smaller number of loops compared to the number of junctions and road segments defined on the two avenues under study, some mathematical approximations were made regarding the positions that had not been accurately measured by the TMC. More specifically, in locations where there was no measuring device, it was considered that the average traffic flow per lane, the average occupancy and the average speed that characterized the road segment was equal to the average of the corresponding traffic parameters that characterized the previous and the next road segment where there was a measuring device and thus a more precise measurement.

$$Q/l_{Sn} = \frac{Q/l_{Sn_{up}} + Q/l_{Sn_{down}}}{2}\left[\frac{Veh}{h}\right] \quad (1)$$

$$O_{Sn} = \frac{O_{Sn_{up}} + O_{Sn_{down}}}{2}\,[\%] \quad (2)$$

$$V_{Sn} = \frac{V_{Sn_{up}} + V_{Sn_{down}}}{2}\,[km/h] \quad (3)$$

Where:

S_n: road segment without a measuring device

Sn_{up}: upstream road segment with a measuring device (before the road segment S_n)

Sn_{down}: downstream road segment with a measuring device (after the road segment S_n)

Q/l_{Sn}: average traffic flow per lane of the road segment S_n

O_{Sn}: average traffic occupancy of the road segment S_n

V_{Sn}: average traffic speed of the road segment S_n

In the case where in a road segment without a measuring device (S_n)

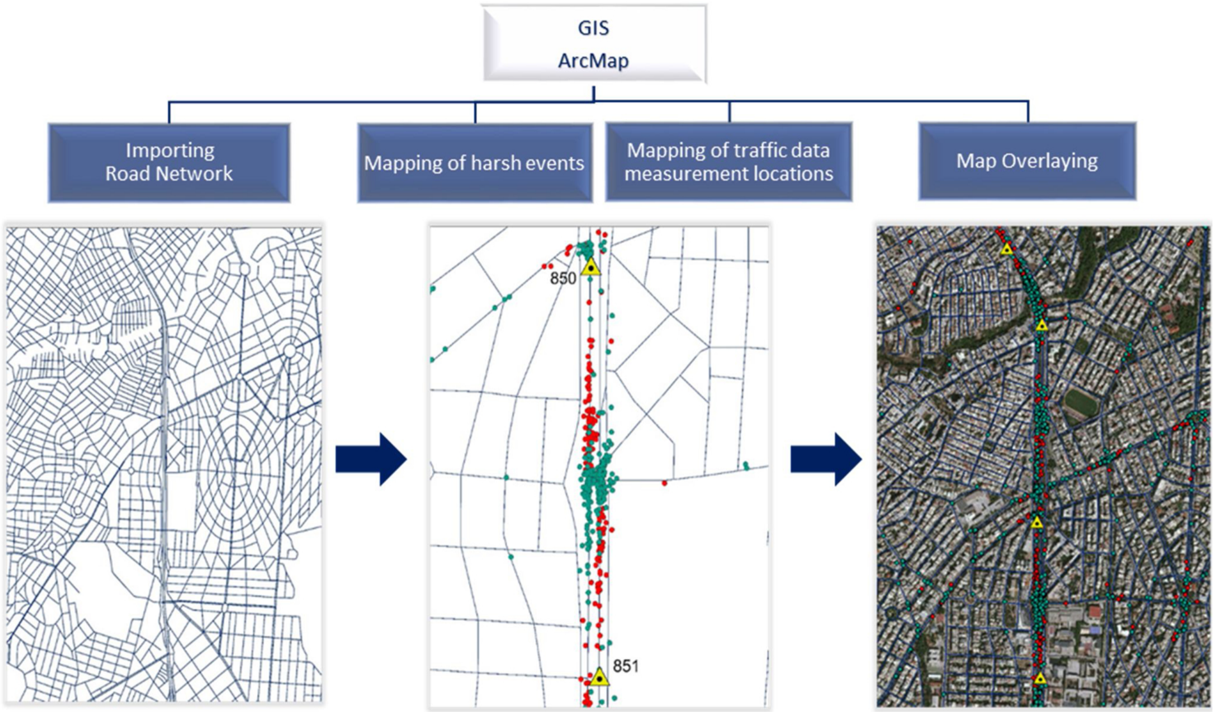

Fig. 2. Visualization of the stages of data processing.

was adding a considerable traffic flow from a road perpendicular to the highway (e.g. Chalkiriou Street in Mesogeion Ave.), a different approach was followed. If S_n was before the perpendicular road, then the traffic parameters were assumed to be equal of those of Sn_{up} only; conversely, if S_n was after the perpendicular road, then the traffic parameters were assumed to be equal of those of Sn_{down} only. This was due to the fact that traffic parameters were typically not available for perpendicular roads, and it was applied in practice only in two instances of the study.

A different process was established for the traffic measurements of junctions. More specific, the average of the traffic measurements of the inbound segments were considered as the overall traffic flow, occupancy and velocity of the junction.

Junction relationships are expressed in mathematical format as

Table 1
Descriptive statistics of collected data.

Variables		Mean	St.Dev.	Min	Max
Road Segments					
Geometric Characteristics	Number of Right Exits & Entrances	1.38	1.53	0	8
	Number of Bus Stops	0.90	0.75	0	3
	Bus Lane	0.46	0.50	0	1
	Sideway	0.46	0.50	0	1
	Segment length [m]	227.78	130.30	20	545
Traffic Characteristics	Traffic Flow /Lane [Veh/h]	554.51	70.77	409.86	693.05
	Velocity [km/h]	54.75	5.17	39.12	63.75
	Occupancy [%]	7.36	1.66	4.65	10.70
Diving Behaviour Characteristics	Harsh Accelerations	26.14	25.37	0	147
	Harsh Brakings	30.72	18.72	2	69
	Range(Speed Diff.) of Harsh Acc.	7.06	3.57	2	16
	Range(Speed Diff.) of Harsh Dec.	7.28	2.97	0.05	13.82
	St.Dev.(Distance) of Harsh Dec.	2.03	1.30	0.35	6.95
Junctions					
Geometric Characteristics	Number of Left Exits	0.82	0.73	0	2
	Number of Left Entrances	0.85	0.83	0	2
	Number of Right Exits	1.03	0.73	0	2
	Number of Right Entrances	1.06	0.75	0	2
	Number of Incoming Lanes	2.06	1.69	0	6
	Number of Outgoing Lanes	1.48	1.54	0	6
	Sideway	0.58	0.50	0	1
Traffic Characteristics	Traffic Flow [Veh/h]	2,588.85	247.71	2,061.39	3,001.90
	Velocity [km/h]	57.20	7.50	46.02	80.24
	Occupancy [%]	7.61	1.52	5.32	9.75
Diving Behaviour Characteristics	Harsh Accelerations	151.12	78.71	2	306
	Harsh Brakings	36.24	32.70	6	164
	St.Dev.(Speed Diff.) of Harsh Acc.	2.52	0.48	1.60	3.80
	MAX(Event Speed) of Harsh Dec.	79.91	19.57	46.28	120.16
	MIN(Distance) of Harsh Dec.	1.36	0.81	0.12	2.32

follows (where k is the total of inbound segments for the junction):

$$Q/l_{jun} = \frac{\sum_{i=1}^{k} Q/l_{inbound}}{k} \left[\frac{Veh}{h}\right] \tag{4}$$

$$O_{jun} = \frac{\sum_{i=1}^{k} O_{inbound}}{k} [\%] \tag{5}$$

$$V_{jun} = \frac{\sum_{i=1}^{k} V_{inbound}}{k} [km/h] \tag{6}$$

In cases where the designed cycles of successive junctions intersected, indicating no intervening road segment, the traffic parameters were considered as equal to the averages of the inbound road segments that entered both junctions, and thus were considered the same for all junctions.

Similar approaches have been followed by past studies conducted with data from the TMC of Athens (Theofilatos and Yannis, 2014; Theofilatos et al., 2018). Since the data is collected historically without any parallel observation, validation is not possible. Any alternative would require the implementation of sophisticated traffic prediction algorithms to impute any missing values, which falls outside the scope of this study.

The data processing was complemented by the introduction of elements related to the geometry of the two expressways under study as they were extracted by visual examination of aerial photography images provided by Google Maps and Google Earth. In the road segments under consideration, variables such as the length of the segment, the number of right exits or entrances, the existence of a sideway etc. were collected. In the junctions under consideration, the number of left and right exits, the number of left and right entrances, the number of incoming and outgoing lanes and the existence of a sideway was determined and inserted in the dataset.

On that similar note, Fig. 3 provides a detail of the examined road environment accompanied by mapping of harsh events. The conditions for defining the limitation and drawing of boundaries for road segments and junctions are also visible.

2.3. Statistical analysis

Following the described data collection and processing, it was decided that multiple linear regression and log-normal regression models would be appropriate for the statistical analyses of driver behaviour. Specifically, regression models were developed to model how parameters of driving behaviour, traffic characteristics and road geometry characteristics influence the frequency of harsh events in a defined spatial unit. In road segments log-normal regression was implemented as it provided a better overall model fit: More accurate predictions were made overall (lower Root Mean Squared Error – RMSE) with log-normal regression, while it prevented negative frequency predictions that might arise from unorthodox independent parameter combinations. In junctions linear regression models were utilized as they provided better model fit and more accurate predictions from log-linear regression models. Predictions of models have been examined in the available dataset, and it was ensured that no negative estimations were made from the linear predictors for the available data.

Linear regression is a widely known, simple technique (and as such we will omit the mathematics behind it) used to model a linear relationship between a continuous dependent variable and one or more independent variables (Washington et al., 2010). In the analysis under consideration, the dependent variables were considered to be continuous due to the large number of events that were observed in the defined spatial units. The log-linear (log-normal) regression was applied as long as it described better the aggressive behaviour in road segments and since all the independent variables in the mathematical models were positive. Both approaches were calibrated using the Ordinary Least Squares method.

A note for the selection of simple linear and log-linear regression models is necessary. The benefits of using Generalised Linear Models (GLMs) for count data have been well-established in the literature (Lord and Mannering, 2010). However, if harsh events are regarded as road segment/junction attributes, especially considering their numbers are significantly higher than crashes for the same temporal period, then linear regression can be explored in order to uncover underlying linear relationships. This is an analogy to economics/econometrics, where linear regression has been implemented to model and predict prices or trading volume (e.g. Gharehchopogh et al., 2013; Yang, 2015). In the opinion of the authors, the discovered strong linear trends shown in the following sections are an incentive to showcase linear/loglinear regression results before delving into GLMs in future research.

To complement the developed models, elasticity analyses were conducted as well. As defined in practice, elasticity analyses allow for the quantification of the response of the dependent variable for a 1% change of an independent continuous variable. When dealing with independent categorical variables, it is meaningful to implement pseudo-elasticities to obtain the incremental changes that are incurred as a result of category changes in the categorical variables (Washington

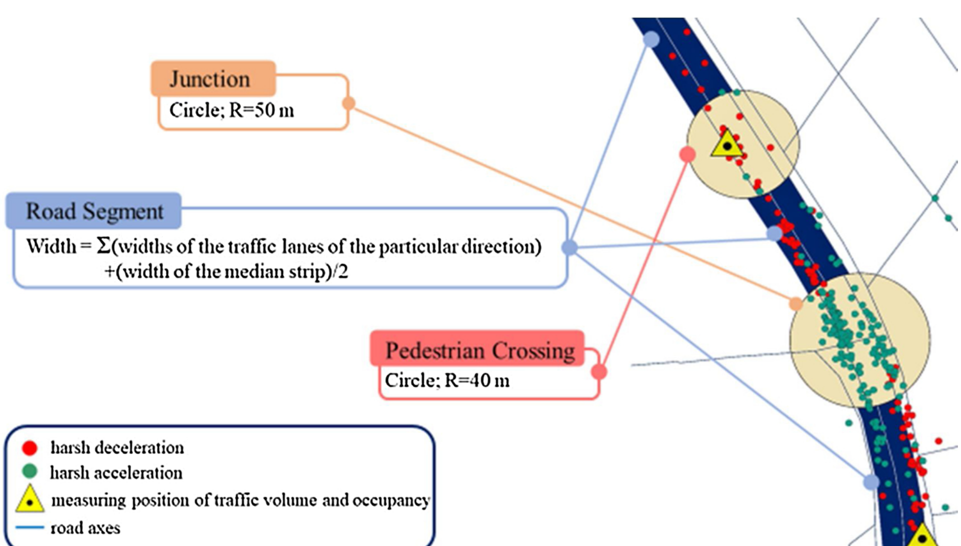

Fig. 3. Example of road environment boundaries and event mapping.

et al., 2010). By using elasticity (and pseudo-elasticity) analyses, the influence of each variable on the number of harsh accelerations or decelerations occurring in a road segment or in a junction was thus quantified.

Following Washington et al. (2010), the elasticity of a dependent variable Y with respect to a continuous independent variable X that has a regression coefficient β can be defined as:

$$e_i = \beta_i \frac{X_i}{Y_i} \approx \frac{\partial X_i}{\partial Y_i} * \frac{X_i}{Y_i} \quad (7)$$

For categorical independent variables, the pseudo-elasticity is defined as per the exponential change:

$$E_{X_{ik}}^{\lambda_i} = \frac{EXP(\beta_k) - 1}{EXP(\beta_k)} \quad (8)$$

The absolute elasticities can be rescaled to fit the range of all independent continuous variables, by setting the lowest value to 1 and adjusting the rest of the variables in proportion with their absolute score. It was decided that it was not appropriate to adjust pseudo-elasticities alongside elasticities as the increases in independent variables are not comparable.

Having established the analysis process, the various model configurations were tested on the data. It should be noted that the final selection of the models was made after several configuration considerations of the many possible combinations of variables, which were documented but are not presented here for brevity. The four final models were evaluated considering the common statistical tests (R^2, *t*-test etc.) but also based on the logical explanation of the results. The correlation of variables was also examined to select the best-fitting mathematical model. In practice, what is expected is the best possible correlation between dependent and independent variables and the zero correlation between independent variables. Those independent variables that showed high correlation, greater than the empirical upper bound of 0.4 were not taken into account in the final behaviour models.

Following that step, models were calibrated using the method of backwards elimination. This process is preferred to the alternatives (such as forward selection or block-wise selection) because the underlying phenomena are not documented well enough to allow for educated guesses of the correct variable mix. Backward elimination is preferred as it provides a better overview of variable importance before the removal of any independent variables.

The analysis was conducted using SPSS Statistics (IBM, 2015). The final model results appear on Table 2 (for road segments) and Table 3 (for junctions) that follow. Variables are considered statistically significant at the typical 95 % level, and they are reported when they are within to up to the 90 % level. For the harsh acceleration model, standard deviation of distance was not found to be statistically significant even at the 90 % level and is thus omitted.

Furthermore, for the developed models, the respective P-P plots were produced, shown in Fig. 4 that follows. The normal probability plots (also known as P-P plots) are used to compare the observed cumulative distribution function (CDF) of the standardized residual to the expected CDF of the normal distribution. It is evident that all of the plots follow closely or very closely the curve, therefore it is assumed that a normal distribution is approximated very adequately by the sample data. Predictions of models have been examined in the available dataset, and it was ensured that no negative estimations were made from the linear predictors for the available data.

3. Discussion of results

3.1. General remarks

An initial result from the descriptive statistical analysis is that the events of harsh accelerations are more frequent at junctions compared to road segments, while the frequency of harsh brakings appears to be marginally higher at junctions than at road segments. At sample level, this may indicate that the available green light time is often violated and many drivers speed up rapidly to evacuate the intersection immediately in orange or even red light duration. When examining the produced models, as is evident from Tables 2 and 3, the variables affecting the frequency of driver harsh events on a statistically significant level in road segments and junctions are several.

The very high adjusted coefficient of determination (R^2) values, especially for the road segments, merit discussion as well. It is not the belief of the authors that they have discovered perfect models that describe harsh events unerringly. High R^2 values alone are not enough to ensure capability of accurate harsh event frequency predictions, however; this much is evident from the sizeable RMSE scores of the models for junctions.

Rather, it is believed that the current research has captured the majority of the parameters that can meaningfully and informatively describe harsh event frequency. This was primarily achieved by taking into consideration three main pillars of road safety: (i) driver behavior, (ii) traffic parameters and (iii) road network/geometric characteristics. It also appears that the relationship is strongly log-linear for road segments and strongly linear in junctions as well. Another probability to keep in mind is that these high values are the product of the dataset, since very homogenous road segments and junctions have been used for its compilation. Ultimately, not all harsh accelerations/decelerations are inherently unsafe. However, by introducing traffic parameters in the models, a large amount of variance in event frequencies has been explained.

The model for harsh deceleration in junctions features a lower R^2 value, apparently due to unobserved factors being present when drivers brake harshly in junctions. This discrepancy – which is not present in the model for harsh decelerations in segments – further highlights not only the inherent difference between the harsh acceleration and harsh

Table 2
Log-linear model results for harsh event frequency in road segments.

Independent Variables	Frequency of harsh events in road segments									
	Harsh accelerations					Harsh decelerations (brakings)				
	Correlation coefficient β_i*	*t* -test value t	p-value p	Absolute elasticity e_i	Relative elasticity e_i*	Correlation coefficient β_i*	*t* -test value t	p-value p	Absolute elasticity e_i	Relative elasticity e_i*
Segment length	0.001	4.224	0.000	0.007	1.138	0.001	2.470	0.017	0.005	1.672
Q/l	0.002	5.974	0.000	0.030	4.858	0.002	4.355	0.000	0.024	8.692
V	−0.011	−2.470	0.017	−0.014	−2.217	−0.010	−1.909	0.063	−0.013	−4.551
Range (Speed Diff.)	0.039	3.477	0.001	0.006	1.000	0.076	4.650	0.000	0.013	4.593
St.Dev. (Distance)						0.058	1.756	0.086	0.003	1.000
Adjusted R^2	0.963					0.954				
RMSE	0.368					0.306				

Table 3
Linear model results for harsh event frequency in junctions.

Independent Variables	Frequency of harsh events in junctions									
	Harsh accelerations					Harsh decelerations (brakings)				
	Correlation coefficient β_i^*	*t* -test value t	p-value p	Absolute elasticity e_i	Relative elasticity e_i^*	Correlation coefficient β_i^*	*t* -test value t	p-value p	Absolute elasticity e_i	Relative elasticity e_i^*
No. Outgoing Lanes	15.520	2.526	0.017	0.104	–					
O	27.857	4.615	0.000	0.014	3.243					
logV^2	−77.010	−3.891	0.001	−0.004	−1.000					
St.Dev. (Speed Diff.)	74.238	3.860	0.001	0.012	2.821					
No. Right Exits						11.436	1.785	0.085	0.338	–
Q						−0.020	−2.135	0.042	0.006	1.000
V^2						0.017	3.339	0.002	0.035	5.511
MAX (Event Speed)						0.491	2.194	0.037	0.011	1.746
MIN (Distance)						−16.160	−2.895	0.007	−0.009	−1.349
Adjusted R^2	0.904					0.723				
RMSE	49.335					23.575				

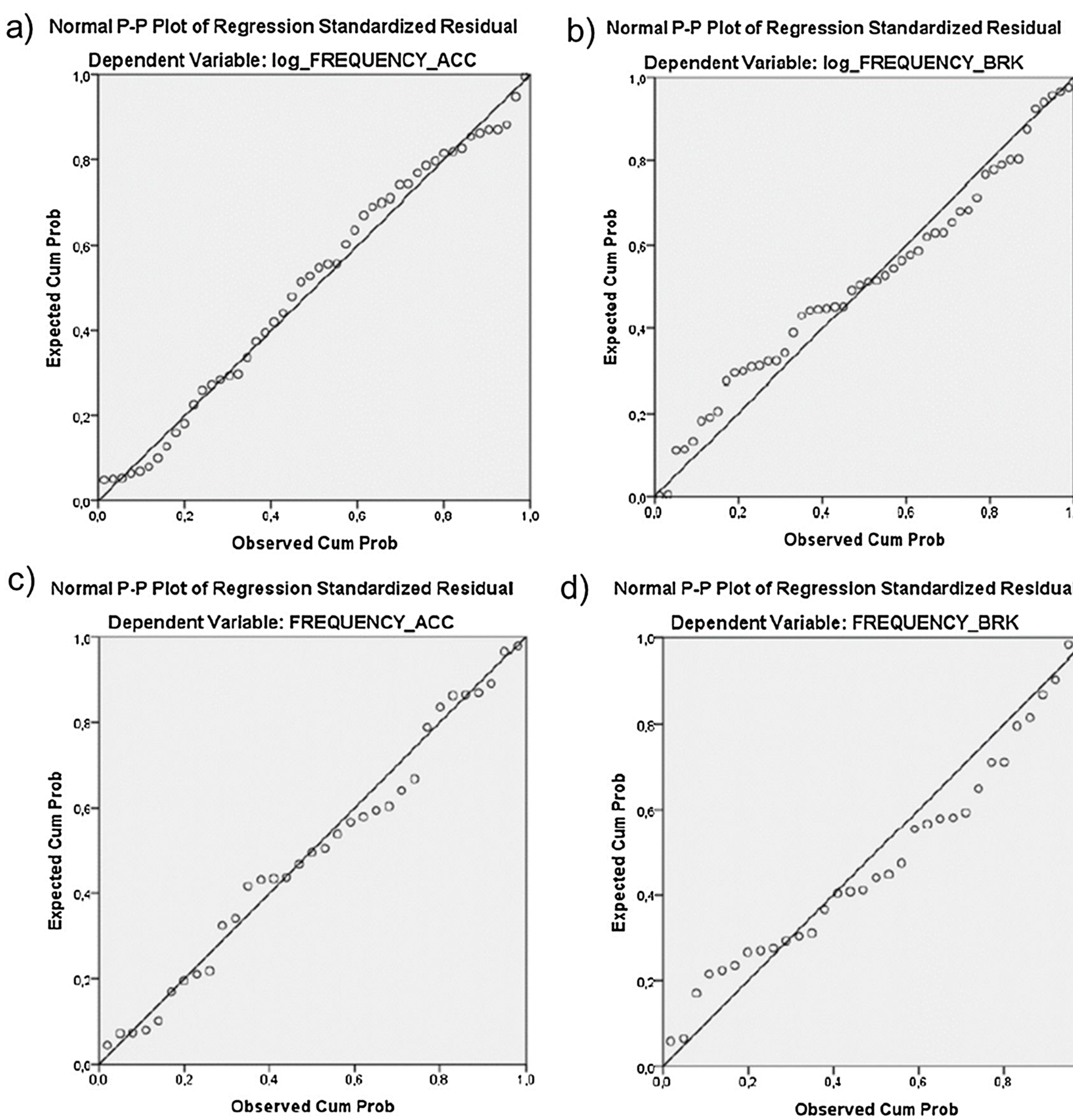

Fig. 4. Produced P-P plots for (a) harsh accelerations in road segments, (b) harsh brakings in road segments, (c) harsh accelerations in junctions, (d) harsh brakings in junctions.

deceleration phenomena, but also the difference between segment and junction environments for harsh braking frequencies.

3.2. Results for road segments

The elasticity analysis for road segments shows that traffic parameters have increased effects than those of road geometry characteristics and driving behavior. In road segments, traffic flow per lane appears as the most critical variable; in particular, an increase of 1% of traffic flow per lane increases the logarithm of the number of harsh accelerations by 0.30 % and the respective value of harsh brakings by 0.24 %. This may be explained by the fact that increased levels of traffic congestion create dynamic obstacles which prevent the driver from selecting their vehicle speeds at will, thus resulting to abrupt driving behavior.

The variables affecting harsh acceleration occurrence in road segments are in order the average speed of traffic, the length of the road segment and the maximum speed difference in two seconds during the event. With regard to the frequency of occurrence of harsh decelerations in road segments second, the most determining factor is the maximum range of differences of speed during the harsh decelerations that occur in the road segment; following in the hierarchy are the average speed of traffic, the length of the road segment and lastly the standard deviation of the distance travelled during the event.

It was observed that as the length of the road segment increases, there is an increase in harsh events of accelerations and decelerations. This may be due the greater number of avenue exits/entrances and the more frequent changes of lanes. Therefore, it could be explained potentially by the fact that while traffic lights are part of the junction areas, their radius of influence extends to downstream and upstream subsections of road segments beyond each cycle defined for each junction. In this way, the aforementioned subsections, have the possibility of a harsh event occurring due to a spillover from the traffic lights ahead.

3.3. Results for junctions

The elasticity analysis for junctions shows that the most important factor affecting harsh accelerations appears to be the average occupancy; an increase of 1% of occupancy increases harsh accelerations by 0.014 %. This result is interpreted because the most aggressive drivers perceive the increase in the saturation of the flow, they try to exploit any available spatial headway. Furthermore, due to speed variation traffic congestion affects driver behavior to a greater extent. It also appears that the average speed of traffic (expressed squared) is the most critical factor in the events of harsh decelerations at junctions, with a 0.5 % increase of speed (resulting to 1% increase of the square of average speed, which is the independent variable used for elasticity analysis) of traffic resulting in a 0.035% increase of harsh brakings. The interpretation is that by increasing average speed of traffic, more events of harsh decelerations are expected, possibly due to the creation of disturbances in traffic waves (shockwaves) or instantaneous queue formation at traffic lights or even from unexpected obstacles.

In the hierarchy of variables that influence the frequency of harsh accelerations in junctions is the maximum difference of speed in two seconds during the event (in the form of standard deviation) and the average speed of traffic. The second most important factor, which affects the number of harsh decelerations that occur in junctions, according to the calculated elasticity, is the maximum speed recorded by the driver speeds during harsh decelerations that took place at the junction. Lastly follows the minimum distance of the harsh deceleration and the average traffic flow.

An increase of the number of outgoing right traffic lanes in a junction causes an increase in harsh accelerations in the junction. Practically this could be explained by the desire of the most aggressive drivers to accelerate since they find more available free space from the increase of the traffic lanes. Another result is that the increase in the number of right exits from the junction causes an increase in the frequency of harsh decelerations which may be justified by the effort of the driver emerging from the junction to adjust the movement speed of his vehicle in a lower speed required to transition to the road network outside of the Avenue. Outgoing vehicles may also encounter other vehicles in the exit lane that drive at a slower speed, and harsh decelerations occur, possibly resulting in a traffic shockwave.

4. Conclusions

The present research aimed to explore the factors affecting the frequencies of harsh accelerations and decelerations (brakings) and their concentrations across junctions and road segments in two urban expressways of Athens. In order to achieve that aim, high resolution smartphone data were utilized for recording the harsh events and driving behavior parameters such as speed difference and event distance from 303 different drivers. This dataset was then complemented by traffic characteristics which were collected through 26 loops, installed by the TMC of Attica Region in specific measuring positions on the two urban expressways under study. Geometrical characteristics of the road network such as lanes were also added by visually examining the area using the online mapping service of Google Maps.

After data collection and subsequent processing, two log-linear regression models were calibrated for road segments and two linear regression models were calibrated for junctions. The models provided valuable insights as a number of affecting factors was determined for harsh event frequency. In short, it was observed that in road segments there is an increase in the number of harsh events if the average traffic flow per lane increases in the respective areas. In junctions as the average occupancy increases, there is an increase in harsh accelerations, and as the average speed increases, more harsh decelerations occur. It appears that traffic characteristics (traffic flow & speed) have the most statistically significant overall impact on the frequency of harsh events compared to road geometry characteristics and driver behavior data.

The important contributions of this paper can be summarized in the following: Firstly, it has been determined that the current research has captured the majority of the parameters that can meaningfully and informatively describe harsh event frequency. This was primarily achieved by taking into consideration three critical pillars of road safety: (i) driver behavior, (ii) traffic parameters and (iii) road network/geometric characteristics. It also appears that the relationship is strongly log-linear for road segments and strongly linear in junctions as well.

Secondly, there seems to be a confirmation of the inherent differences of the nature of events of harsh accelerations and harsh decelerations. Modelling results indicate that the frequencies of these events are not described by identical mixes of variables, and are also influenced by different factors in segments and in junctions.

The results of this study may be transferred to similar areas outside the research area. However, prior to any generalization, necessary adjustments should be made for possible variations in the road environment and traffic. For instance, an analogous study should be conducted for motorways or rural roads that have fundamentally different characteristics than urban expressways in order to obtain more accurate results for these road environments. Alternative count models such as GLMs with known merits for similar research or machine learning methods should also be investigated. Initial Poisson loglinear model applications have shown that the discovered relationships are retained in significance and sign (positive or negative influences). Intuitively, several crash frequency methods found in the rich road safety literature have the potential to be applied in harsh event investigation, and the respective findings will augment and expand the knowledge obtained from strictly analyzing crashes.

Furthermore, a very promising direction for future research would be the investigation of crash numbers and locations in the same research areas. It would be fruitful to test correlations of crash frequencies with some of the variables that have been identified in the present study, and to explore any correlations between harsh event frequencies and crash frequencies as well. However, it is worth noting that this endeavor requires significant updates in crash data collection procedures in Greece, as crash locations tend to be very imprecise compared to high-resolution smartphone data. A similar conundrum rises when weather data are brought into consideration. The inclusion of weather data in the present context would be quite interesting, as there can be considered to be related to crashes from research (Theofilatos and Yannis, 2014), and to harsh events from observation and experience. However, the high resolution smartphone data utilized in the study would be best paired with comparably high resolution weather data, which are at present not readily available. Therefore, further research is needed to create a proper smartphone naturalistic driving data and weather data merging scheme, which will yield usable results towards this direction.

CRediT authorship contribution statement

Virginia Petraki: Data curation, Visualization, Formal analysis, Writing - original draft, Writing - review & editing. **Apostolos Ziakopoulos:** Methodology, Resources, Writing - original draft, Writing - review & editing. **George Yannis:** Methodology, Conceptualization, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is co-financed by Greece and the European Union (European Social Fund - ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project “Strengthening Human Resources Research Potential via Doctorate Research” (MIS-5000432), implemented by the State Scholarships Foundation (IKY).

The authors would like to thank OSeven Telematics, London, UK for providing all necessary driver behavior data exploited to accomplish this study. The authors also like to thank the Traffic Management Centre of Attica for providing all traffic parameter data exploited to accomplish this study.

References

Abdel-Aty, M., Wang, X., 2006. Crash estimation at signalized intersections along corridors: analyzing spatial effect and identifying significant factors. Transp. Res. Rec. 1953 (1), 98–111.
Abdel-Aty, M., Lee, J., Siddiqui, C., Choi, K., 2013. Geographical unit based analysis in the context of transportation safety planning. Transp. Res. Part A Policy Pract. 49 (March), 62–75.
Alessandrini, A., Cattivera, A., Filippi, F., Ortenzi, F., 2012. Driving style influence on car CO2 emissions. In: Proceedings of the 20th International Emission Inventory Conference - "Emission Inventories - Meeting the Challenges Posed by Emerging Global. National, and Regional and Local Air Quality Issues". Tampa, Florida. Aug.
ArcGIS 10.3 GIS Software, 2020. ArcGIS 10.3 GIS Software. ESRI, New York.
ArcMap 10.3 GIS Software, 2020. ArcMap 10.3 GIS Software. ESRI, New York.
Bonsall, P., Liu, R., Young, W., 2005. Modelling safety-related driving behaviour—impact of parameter values. Transp. Res. Part A Policy Pract. 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. Proc. Natl. Acad. Sci. 113 (March 10), 2636–2641.
ELSTAT, 2020. Hellenic Statistical Authority. www.statistics.gr/en/home/.
Essa, M., Sayed, T., 2018. Traffic conflict models to evaluate the safety of signalized intersections at the cycle level. Transp. Res. Part C Emerg. Technol. 89, 289–302.
European Commission, 2018. Road Safety Statistics: What Is Behind the Figures? April 2019. europa.eu/rapid/press-release_MEMO-19-1990_en.htm?fbclid=IwAR1aBc2E1Z4d49X8RAc6cEcZNHe0fz2HH3iCUwX4pwDctp5IhddpOx-m0XU. .
European Commission, 2019. Annual Accident Report. European Commission, Directorate General for Transport. June 2018.. . ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/statistics/dacota/asr2018.pdf.
Gharehchopogh, F.S., Bonab, T.H., Khaze, S.R., 2013. A linear regression approach to prediction of stock market trading volume: a case study. Int. J. Manag. Value Supply Chain. 4 (3), 25.
Gündüz, G., Yaman, Ç., Peker, A.U., Acarman, T., 2017. Prediction of risk generated by different driving patterns and their conflict redistribution. Ieee Trans. Intell. Veh. 3 (1), 71–80.
Guo, Y., Essa, M., Sayed, T., Haque, M.M., Washington, S., 2019. A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in Australia. Transp. Res. Part C Emerg. Technol. 101, 96–110.
Hamdar, S.H., Qin, L., Talebpour, A., 2016. Weather and road geometry impact on longitudinal driving behavior: exploratory analysis using an empirically supported acceleration modeling framework. Transp. Res. Part C Emerg. Technol. 67 (June), 193–213.
Hanowski, R.J., Perez, M.A., Dingus, T.A., 2005. Driver distraction in long-haul truck drivers. Transp. Res. Part F Traffic Psychol. Behav. 8 (6), 441–458.
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. Accid. Anal. Prev. 38 (January 1), 185–191.
Huang, F., Liu, P., Yu, H., Wang, W., 2013. Identifying if VISSIM simulation model and SSAM provide reasonable estimates for field measured traffic conflicts at signalized intersections. Accid. Anal. Prev. 50, 1014–1024.
IBM, 2015. IBM SPSS Statistics for Windows, Version 23.0. IBM Corp., Armonk, NY.
Jansen, R.J., Wesseling, 2018. SWOV. Harsh braking by truck drivers: a comparison of thresholds and driving contexts using naturalistic driving data. In Proceedings of the 6th Humanist Conference.
Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J., 2009. Comparing Real-world Behaviors of Drivers With High Versus Low Rates of Crashes and Near-crashes. NHTSA Final Letter Report, Washington, DC, USA Feb.
Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transp. Res. Part A Policy Pract. 44 (5), 291–305.
Mantouka, E.G., Barmpounakis, E.N., Vlahogianni, E.I., 2018. Mobile sensing and machine learning for identifying driving safety profiles. In: Presented at 97th Annual Meeting of the Transportation Research Board. Washington, D.C..
Olson, R.L., Hanowski, R.J., Hickman, J.S., Bocanegra, J., 2009. Driver Distraction in Commercial Vehicle Operations (No. FMCSA-RRT-09-042). Federal Motor Carrier Safety Administration, United States.
Paefgen, J., Kehr, F., Zhai, Y., Michahelles, F., 2012. Driving behavior analysis with smartphones: insights from a controlled Field study. In: Proceedings of the 11th International Conference on mobile and Ubiquitous Multimedia. ACM. pp. 36 December.
Paefgen, J., Staake, T., Fleisch, E., 2014. Multivariate exposure modeling of accident risk: insights from Pay-as-you-drive insurance data. Transp. Res. Part A Policy Pract. 61, 27–40.
Papadimitriou, E., Tselentis, D.I., Yannis, G., 2018. Analysis of driving behavior characteristics based on smartphone data. Proceedings of 7th Transport Research Arena TRA 16–19.
Papadimitriou, E., Argyropoulou, A., Tselentis, D.I., Yannis, G., 2019. Analysis of driver behaviour through smartphone data: the case of mobile phone use while driving. Saf. Sci.
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. Transp. Res. Part F Traffic Psychol. Behav. 6 (December 4), 275–287.
Shaout, A.K., Bodenmiller, A.E., 2011. A Mobile application for monitoring inefficient and unsafe driving behaviour. International Arab Conference on Information Technology (ACIT) 1–8 Dec 11.
Singh, S., 2015. Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey NHTSA Report (No. DOT HS 812 115).
Singh, S., 2018. Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey. (Traffic Safety Facts Crash Stats. Report No. DOT HS 812 506). National Highway Traffic Safety Administration., Washington, DC.
Stavrakaki, A.M., Tselentis, I.D., Barmpounakis, E.N., Vlahogianni, I.E., Yannis, G., 2019. How Much driving data Do We need to assess driver behavior? In: Presented at 98th Annual Meeting of the Transportation Research Board. Washington, D.C..
Stephens, A.N., Groeger, J.A., 2009. Situational specificity of trait influences on drivers’ evaluations and driving behaviour. Transp. Res. Part F Traffic Psychol. Behav. 12 (1), 29–39.
Theofilatos, A., 2017. Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials. J. Safety Res. 61, 9–21.
Theofilatos, A., Yannis, G., 2014. A review of the effect of traffic and weather characteristics on road safety. Accid. Anal. Prev. 72, 244–256.
Theofilatos, A., Yannis, G., Vlahogianni, E.I., Golias, J.C., 2017. Modeling the effect of traffic regimes on safety of urban arterials: the case study of athens. J. Traffic Transp. Eng. 4 (3), 240–251.
Theofilatos, A., Yannis, G., Antoniou, C., Chaziris, A., Sermpis, D., 2018. Time series and support vector machines to predict powered-two-wheeler accident risk and accident type propensity: A combined approach. J. Transp. Saf. Secur. 10 (5), 471–490.
Toledo, T., Oren, M., Tsippy, L., 2008. In-vehicle data recorders for monitoring and feedback on drivers’ behavior. Transp. Res. Part C Emerg. Technol. 16 (June 3), 320–331.
Troncoso, C., Danezis, G., Kosta, E., Balasch, J., Preneel, B., 2010. Pripayd: privacy-

friendly pay-as-you-drive insurance. IEEE Trans. Dependable Secure Comput. 8 (December 5), 742–755.
Tselentis, D.I., Yannis, G., Vlahogianni, E.I., 2017. Innovative motor insurance schemes: a review of current practices and emerging challenges. Accid. Anal. Prev. 98, 139–148.
Tselentis, D.I., Vlahogianni, E.I., Yannis, G., 2018. Comparative evaluation of driving efficiency using smartphone data. In: Presented at 97th Annual Meeting of the Transportation Research Board. Washington, D.C.
Tselentis, D., Vlahogianni, E., Yannis, G., 2019. Investigating the temporal evolution of driving safety efficiency using data collected from smartphone sensors. In: Presented at 98th Annual Meeting of the Transportation Research Board. Washington, D.C..
Vaiana, R., Iuele, T., Astarita, V., Caruso, M.V., Tassitani, A., Zaffino, C., Giofrè, V.P., 2014. Driving behavior and traffic safety: an acceleration-based safety evaluation procedure for smartphones. Mod. Appl. Sci. 8 (February 1), 88.
Vlahogianni, E.I., Barmpounakis, E.N., 2017. Driving analytics using smartphones: algorithms, comparisons and challenges. Transport. Res. C: Emerg. Technol 79 (June), 196–206.
Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2010. Satistical and Econometrics Methods for Transportation Data Analysis. Chapman and Hall/CRC.
World Health Organization, 2018. Global Satus Report on Road Safety.
Wouters, P.I., Bos, J.M., 2000. Traffic accident reduction by monitoring driver behaviour with in-car data recorders. Accid. Anal. Prev. 32 (5), 643–650.
Yamakado, M., Takahashi, J., Saito, S., Abe, M., 2009. G-vectoring: new vehicle dynamics control technology for safe driving. Industrial systems. Hitachi Review 58 (December 7), 347.
Yang, Y., 2015. Development of the regional freight transportation demand prediction models based on the regression analysis methods. Neurocomputing 158, 42–47.
Yannis, G., Papadimitriou, E., Folla, K., 2014a. Effect of GDP changes on road traffic fatalities. Saf. Sci. 63, 42–49.
Yannis, G., Theofilatos, A., Ziakopoulos, A., Chaziris, A., 2014b. Investigation of road accident severity. Traffic Eng. Control 55 (1).
Yannis, G., Tselentis, D., Vlahogianni, E., Argyropoulou, A., 2017. Monitoring distraction through smartphone naturalistic driving experiment. In: Proceedings of the 6th International Naturalistic Driving Research Symposium (NDRS). The Hague. 8-9 June.
Ziakopoulos, A., Yannis, G., 2020a. A review of spatial approaches in road safety. Accid. Anal. Prev. 105323.
Ziakopoulos, A., Yannis, G., 2020b. Meta-regressions of exposure parameters used in spatial road safety analyses. Adv. Transp. Stud.: Int. J. 51.
Ziakopoulos, A., Tselentis, D., Kontaxi, A., Yannis, G., 2020. A critical overview of driver recording tools. Journal of Safety Research.s.
Zohar, D., Huang, Y.H., Lee, J., Robertson, M., 2014. A mediation model linking dispatcher leadership and work ownership with safety climate as predictors of truck driver safety performance. Accid. Anal. Prev. 62, 17–25.