Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2022) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

Transport Research Arena (TRA) Conference

# Critical factors affecting mobile phone use while driving through the exploitation of smartphone sensor data

Sofia Akritidou[a], Armira Kontaxi[a]*, Apostolos Ziakopoulos[a], George Yannis[a]

[a]*National Technical University of Athens, 5 Heroon Polytechniou St., GR-15773, Athens, Greece*

**Abstract**

The aim of the present research is to model the critical factors of mobile phone use on driver behavior through the exploitation of data from smartphone sensors. Data collected from 100 drivers who participated at a naturalistic driving experiment for six months are analyzed, combined with respective questionnaire answers. Using GLM Poisson regression models, the correlation between driving characteristics recorded by smartphone sensors and the percentage of mobile use (converted to integers) while driving was investigated. Four statistical regression models forecasting the percentage of mobile phone use while driving were developed: one overall model and one per road type (urban, rural, highway). Results indicate that parameters affecting the use of mobile phone while driving are (i) the percentage of driving duration with speed above the speed limit (ii) driving distance, (iii) average deceleration, and (iv) average speed. Across the four models, average deceleration had the most consistent impact, appearing to have a statistically significant negative correlation with mobile use integer values.

*Keywords:* road safety;driver behavior; mobile phone use; smartphone data; GLM Poisson regression models

## 1. Introduction

Mobile phone use and engagement while driving is one of the most critical factors affecting road safety levels and one of the leading cause of road crashes (Haque & Washington, 2015). Drivers who use a mobile phone while driving exhibit impairment in their driving performance (higher speed variations, difficulties in maintaining vehicle lateral position, etc.) due to high levels of workload. As a result, increased reaction times are being observed in case of an incident or event and therefore high probability of being involved in a crash. Despite significant improvement of road safety measures by countries and road management authorities, the number of related fatalities and injuries remains plateaued at high levels. Therefore, increased effort is required to identify crucial factors affecting mobile phone use

* Corresponding author. Tel.: +30-210-772-1575; *E-mail address:* akontaxi@mail.ntua.gr


Peer-review under responsibility of the scientific committee of the Transport Research Arena (TRA) Conference

while driving by employing evidence-based approaches (Choudhary & Velaga, 2017; Papadimitriou et al., 2019; Papantoniou et al., 2020).

Simultaneously, mobile phones have advanced at a very rapid pace in recent years. Shifting from a basic communication device to a programmable, adaptable, affordable and intuitive interconnected platform, the replacement of legacy mobile phones with smartphone devices is very high in industrialized societies. Several applications (apps) have been created fulfilling various niche roles for transport activities and road safety evaluations. For instance, smartphones have been used to develop applications conducting specialized driver monitoring, and can be used to measure harsh driving events and analyse them macroscopically (Petraki et al., 2020; Kontaxi et al., 2021) or microscopically (Ziakopoulos, 2021). Another example is a mobile-to-mobility app which has been developed and tested in Cosenza, Italy, aiming to collect and disseminate crowdsourced data and information concerning road and traffic functional characteristics, and interviewees reported different willingness-to-use the app for different network characteristics and conditions, though the influence of the number of isolated incidents to the measured willingness was not quantified (Cardamone et al., 2014).

Overall, smartphones can be considered as an attractive data collection approach by researchers for driver monitoring due to the fact that they are extremely widespread, allowing for continuous, rapid data collection and have flexible capabilities through programing, while individual data collection is quite low-cost and seamless once there is the required digital infrastructure is operational. On the negative side, smartphone-based data collection has steep data storage and data analysis requirements, while the upfront costs can be considerable as well before the scheme is operational (Ziakopoulos, et al., 2020).

Based on the aforementioned, the aim of this study is to examine and model the critical factors that affect mobile phone use while driving through the exploitation of data from smartphone sensors. More specifically, the authors investigate the effect of road type (urban, rural and highway) in the percentage of mobile use while driving as well as the correlation with some driving characteristic such as driving distance. To achieve this objective, data collected from more than 200 drivers who participated in a naturalistic driving experiment for six months via the OSeven smartphone application are analysed, combined with the answers of 100 drivers to a questionnaire based on their demographic and driving behaviour characteristics.

## 2. Data collection

This study exploits data from a 200-driver naturalistic experiment in which drivers conducted their daily trips normally while having the application of OSeven (www.oseven.io) installed in their smartphone devices. Within a 6-month timespan (from July to December 2019), a total of 49,019 trips were undertaken. The solid integration platform for collecting, transferring raw data and recognizing the driving behaviour metrics via ML algorithms is also developed by OSeven. The data flow system every time a new trip is recorded by the application is clearly presented in Figure 1.


Fig. 1. The OSeven data flow system.

In every trip a driver completes, a large amount of data is recorded, transmitted through WiFi or cellular network and critical information such as metrics, features, highlights and driving score is produced in order to evaluate driving profile and performance. More precisely, the available naturalistic driving data include, indicatively, trip distance and duration, road type where the vehicle is driving, position, vehicle speed, speeding (duration and amount of limit exceedance), harsh braking/acceleration events and duration in which the smartphone is actively used by the driver while moving; for a fuller description, the reader is referred to Papadimitriou et al. (2019). The latest parameter is used to determine driver distraction via smartphone use and, as a percentage, serves as the dependent variable of the present analysis.

An indicative visualization of the distribution of the percentage of driving mobile phone use (Mbu) and the respective trend lines is shown on Figure 1, examined by (a) the speeding percentage and (b) the total distance of the trip. The charts reveal that:

- As the use of the mobile phone increases, the rate of exceeding the speed limits almost stabilizes up to about 75%. From 75% -100% of the percentage of mobile phone use, there is a small increase in the rate of exceeding the speed limit.
- As the mobile phone usage rate increases, the total driving distance remains constant.

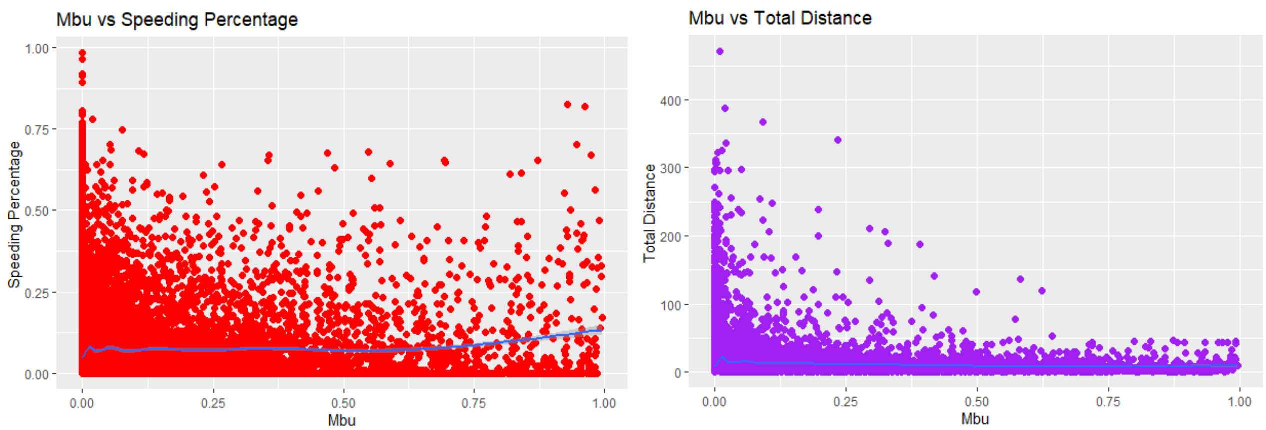


Fig. 2. Mobile phone use by (a) speeding percentage; (b) total trip distance.

Complementary to the naturalistic driving data, a questionnaire was issued and collected from participants. The questionnaire featured detailed questions about driving experience, habits and behavior, as well as demographic characteristics. The questionnaire had numerous questions that are omitted for brevity reasons. As per example, Figure 3 illustrates the number of drivers in regards to their (a) age and (b) gender distribution and the self-declaration of the frequency of mobile phone use while driving.

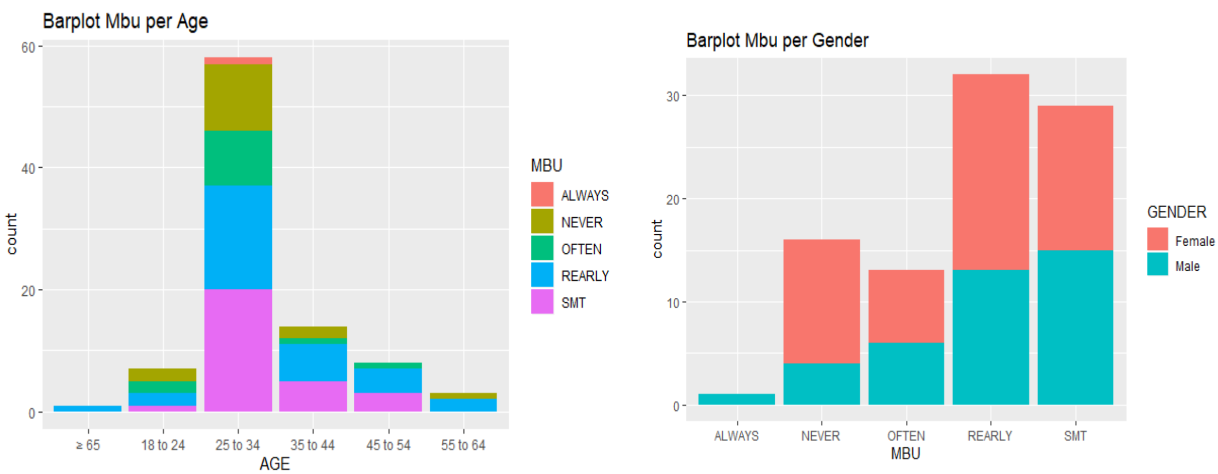


Fig. 3. Drivers' distribution by (a) age and (b) gender and self-declared mobile phone use while driving.

The above charts illustrate that:

- With respect to the age and use of the mobile phone, it is found that the largest number of drivers who reportedly make continuous (ALWAYS) or sometimes (SMT) use of the mobile phone are aged 25 to 34 years.
- Regarding mobile phone use and gender, male drivers declare to use more frequently their mobile phone while driving, according to the questionnaire answers. On the contrary, women state that they rarely use their mobile phones while driving.

## 3. Results

From the aforementioned datasets, data from drivers that had both answered the questionnaire in full and completed successfully more than 10 trips was utilized; this corresponded to 89 remaining drivers. Subsequently, Poisson models were developed in order to examine whether driving characteristics recorded by smartphone sensors and driver characteristics from the questionnaire affect and can therefore predict the percentage of mobile use while driving. Precisely, after transforming the percentage of mobile use per trip to an integer, Generalized Linear Models (GLMs) were implemented with a Poisson data distribution. GLMs are known to be better used when dealing with frequency (count) data (Lord & Mannering, 2010).

The general form of the GLM models the log odds via a linear predictor. Following McCulloch (2003), if y is the observed speeding percentage per trip i, and λ is the expected speeding percentage to be predicted, then the model is specified as:

$$y_i \sim Poisson\ (\lambda_i) \tag{1}$$

And the linear predictor is:

$$\log(\lambda_i) = \beta_o + \beta_n + x_n + \varepsilon \tag{2}$$

Where $\beta$ are the fixed-effect parameters (constant and coefficients) for *n* independent variables, and $\varepsilon$ is the error term.

Variable multicollinearity was investigated, and variable pairs that had more than 0.5 Pearson scores were mutually excluded from the model formulation process, which was executed with the backwards elimination technique. Ultimately, four statistical models for mobile use percentage were developed: one overall model and three models for each road type (urban, rural, highway). All analyses were conducted using R-studio (R Core Team, 2019). Table 1 provides a description of the variables selected.

Table 1. Description of the variables used in the analyses

| Variables | Explanation |
|---|---|
| speeding_perc | percentage of speeding time while driving (%) |
| distance | trip distance (km) |
| dec_avg | average deceleration per trip ($m/s^2$) |
| speed_avg | average driving speed |
| km_avg_day | self-declaration of average travelled distance per day (km) |
| sp_familiarity | drivers' familiarity on smartphones (more than average=1, less than average =0) |
| self-declared mbu | self-declaration of mobile use while driving (never=1, rarely=2, sometimes=3, often=4, always=5) |
| AIC | Akaike information criterion |
| McFadden | McFadden's pseudo $R^2$ |

The final models are presented in Table2; overall model and urban, rural and highway models, respectively. A number of interesting observations emerge from the factors of the above mathematical models: in all four regression models the frequency of deceleration is associated with reduced mobile use while driving. Furthermore, in three out of four models the speeding percentage, the total distance of the trip and the self-reported familiarity of the driver with the use of the mobile phone are associated with increased mobile use while driving. The self-reported frequency of the mobile use by drivers is also found statistically significant. An additional interesting finding is that the highway model contains different variables from the other three models, indicating a different driving behaviour pattern. More precisely, the model results could be further interpreted, computing the relative risk ratio of every independent variable and thus measuring the increase in the probability of mobile phone use while driving.

Table 2. Poisson models for the percentage of mobile phone use while driving

| Parameter | Overall model | | Urban model | | Rural model | | Highway model | |
|---|---|---|---|---|---|---|---|---|
| | B | P-value | B | P-Value | B | P-value | B | P-value |
| intercept | -1.223 | 0.097 | 0.963 | 0.025 | 0.411 | 0.004 | -1.941 | <0.001 |
| speeding_perc | 4.112 | <0.001 | 2.841 | <0.001 | 7.358 | <0.001 | - | - |
| distance | 0.032 | <0.001 | -0.122 | <0.001 | - | - | 7.932 | <0.001 |
| dec_avg | -1.383 | <0.001 | -0.625 | 0.005 | -1.182 | 0.008 | -0.075 | <0.001 |
| speed_avg | - | - | - | - | - | - | -0.820 | <0.001 |
| km_avg_day | - | - | -0.355 | 0.046 | -0.497 | 0.045 | - | - |
| sp_familiarity | 1.172 | <0.001 | 1.070 | 0.017 | 1.376 | <0.001 | - | - |
| self-declared mbu never | -2.433 | <0.001 | -2.532 | <0.001 | -2.556 | <0.001 | - | - |
| self-declared mbu often | -0.729 | <0.001 | -0.753 | <0.001 | -0.558 | <0.001 | - | - |
| self-declared mbu rarely | -1.697 | <0.001 | -1.769 | <0.001 | -1.746 | <0.001 | - | - |
| self-declared mbu smt | -1.178 | <0.001 | -1.151 | <0.001 | -1.089 | <0.001 | - | - |
| AIC | 527.08 | | 527.94 | | 438.61 | | 148.93 | |
| McFadden pseudo $R^2$ | 0.258 | | 0.253 | | 0.321 | | 0.241 | |

First, the variables collected from the smartphone application are going to be examined. Specifically, the driving behaviour indicator of speeding (speeding_perc) while driving seem to increase the odds of mobile phone use during the trip; the effect is found to be higher in rural areas, less in the overall model and the least in urban areas, as the respective odds ratios exp (B) are 1568.4, 62.0, and 17.1. A probable explanation might be that the driver who does not comply with the speed limits, does not comply with the forbiddance of the use of the mobile phone while driving. This implies a risky driving profile. An interesting finding is the fact that the mentioned variable is not statistically important in the highway model; on the contrary, we can observe that the variable of mean speed, which is found statistically significant, appears to correlate negatively with the use of mobile phone, i.e. the increase of mean speed decreases the odds of mobile phone use during the trip. Taking into consideration the highway environment, the most likely explanation is that the higher the driving speeds, the more difficult the driving task and the more concentration is required by the driver. Therefore the mobile phone usage rate decreases.

Additionally, the variable of mean deceleration (dec_avg) is statistically significant in all the four models and negatively correlated with the use of mobile phone while driving. More specifically, the effect is found to be higher in highways, less in urban areas and the least in rural areas and the overall model, as the respective odds ratios exp (B) are 0.9, 0.5, and 0.3.

Finally, the exposure metric of the driven distance per trip seems to increase the odds of mobile phone use while driving both in the overall and the highway model by a factor of 1.4 and 79.3, respectively. The high impact of the variable in the highway model can be explained by the fact that long distances in highways may be conceived as repetitious, resulting in the increased probability of using the mobile phone in terms of talking/texting/surfing the internet. However, in the urban environment, it seems that the more the travelled distance the smaller the probability of using the mobile phone. This may be explained by the fact that longer distances in an urban environment mean a greater amount of information (e.g. manoeuvres, pedestrians, signalling, etc.) that the driver has to process, thus reducing mobile phone usage.

Regarding the variables collected from the questionnaire filled in by the participants, three are found to be statistically significant, namely the declared travelled distance per day, the participants' familiarity on using a smartphone and the self-declaration of mobile use while driving. It is interesting to observe that drivers who declare to be quite familiar with the use of smartphone applications, are prone to use the mobile phone while driving, revealing that the more familiar the driver is with the mobile phone, the more confident they feel about using it while driving. More precisely, the effect is found to be higher in rural areas, less in the overall model and the least in urban areas, as the respective odds ratios exp (B) are 4.0, 3.3 and 3.0.

A for the variable of the self-declaration of average travelled distance per day, it is found that as the daily travelled distance increases, especially for distances longer than 15 km, the use of the mobile phone decreases by a factor of 1.4 in urban areas and by a factor of 1.6 in rural areas. The explanation might be that the driving environment is more complex, so the use of mobile phones is reduced.

Last but not least, regarding the self-declared mbu variable, i.e. the self-declared use of the mobile phone while driving, it seems that the drivers' answers to the questionnaire are true as the less often they answer that they use the mobile phone, the more the actual use of the mobile phone (as monitored by the OSeven application) decreases. While the more often the drivers declare to use the mobile phone while driving, the more increases the actual use of the mobile phone.

## 4. Discussion

The present study aimed to model the influence of critical factors of mobile phone use on driver behavior through the exploitation of data from smartphone sensors. Data collected from 100 drivers who participated at a naturalistic driving experiment for six months are analyzed, combined with respective questionnaire answers. Using Poisson GLM models, the correlation between driving characteristics recorded by smartphone sensors and the percentage of mobile use while driving was investigated. Four statistical regression models forecasting the percentage of mobile using while driving were developed: one overall model and one per road type (urban, rural, highway).

Results indicate that parameters affecting the use of mobile phone while driving are (i) the percentage of driving duration with speed above the speed limit (ii) driving distance, (iii) average deceleration, and (iv) average speed. Additionally, the questionnaire variables found to be statistically significant are: (i) declared travelled distance per day, (ii) the participants' familiarity on using a smartphone and (iii) the self-declaration of mobile use while driving.

An interesting finding of the present study is that across the four models, average deceleration had the most consistent impact, appearing to have a statistically significant negative correlation with mobile use integer values. Furthermore, another significant outcome is worthy of exploration in future research is the positive correlation of the percentage of speeding time and mobile phone use, indicating that drivers who don't comply with the speed limits are those who don't comply with the banning of mobile phone use while driving.

Future research could focus on the analysis of other driving behavior parameters identified by the road safety literature as risk factors (e.g. exceeding speed limit, harsh events frequency) and their effect on driving performance and road safety. Furthermore, analyses per age, crash history, self–assessment, driving experience and more demographic characteristics could be undertaken in order to capture any particular trends found in the categories of these parameters. Finally, the results of the analysis could allow for recommendations regarding risk factors that appear to be critical for safe driving, exploited by stakeholders including policy makers and industry.

## Acknowledgements

This research is co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program "Competitiveness, Entrepreneurship and Innovation", under the call "RESEARCH-CREATE-INNOVATE" (project code: T2EΔK-04388).

## References

Cardamone, A. S., Eboli, L., Forciniti, C., & Mazzulla, G. (2014). Willingness to use mobile application for smartphone for improving road safety. International journal of injury control and safety promotion, 23(2), 155-169.

Choudhary, P., & Velaga, N. R. (2017). Mobile phone use during driving: Effects on speed and effectiveness of driver compensatory behaviour. Accident Analysis & Prevention, 106, 370-378.

Haque, M. M., & Washington, S. (2015). The impact of mobile phone distraction on the braking behaviour of young drivers: a hazard-based duration model. Transportation research part C: emerging technologies, 50, 13-27.

Kontaxi, A., Ziakopoulos, A., & Yannis, G. (2021). Trip characteristics impact on the frequency of harsh events recorded via smartphone sensors. IATSS Research, 45(4), 574-583.

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transportation research part A: policy and practice, 44(5), 291-305.

McCulloch, C. E. (2003). Generalized linear mixed models. In NSF-CBMS regional conference series in probability and statistics 7. Institute of Mathematical Statistics and the American Statistical Association (IMS), Beachwood, OH. MR1993816. (pp. i-84).

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. Safety Science, 119, 91-97.

Papantoniou, P., Kontaxi, A., Yannis, G., & Fortsakis, P. (2020, June). Investigating the correlation of mobile phone use with trip characteristics recorded through smartphone sensors. In Conference on Sustainable Urban Mobility (pp. 935-944). Springer, Cham.

Petraki, V., Ziakopoulos, A., & Yannis, G. (2020). Combined impact of road and traffic characteristic on driver behavior using smartphone sensor data. Accident Analysis & Prevention, 144, 105657.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Ziakopoulos, A. (2021). Spatial analysis of harsh driving behavior events in urban networks using high-resolution smartphone and geometric data. Accident Analysis & Prevention, 157, 106189.

Ziakopoulos, A., Tselentis, D., Kontaxi, A., & Yannis, G. (2020). A critical overview of driver recording tools. Journal of safety research, 72, 203-212.