Transportation Research Part C 109 (2019) 343–357

Contents lists available at ScienceDirect

Transportation Research Part C

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

Driving safety efficiency benchmarking using smartphone data

Dimitrios I. Tselentis*, Eleni I. Vlahogianni, George Yannis

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

ARTICLE INFO

Keywords:
Driving behavior
Driving safety efficiency
Data envelopment analysis
Smartphone data

ABSTRACT

This paper aims to provide a methodological framework for the comparative evaluation of driving safety efficiency based on Data Envelopment Analysis (DEA). The analysis considers each driver as a Decision Making Unit (DMU) and aims to provide a relative safety efficiency measure to compare different drivers based on their driving performance. The last is defined based on a set of driving analytics (e.g. distance travelled, speed, accelerations, braking, cornering and smartphone usage) collected using an innovative data collection scheme, which is based on the continuous recording of driving behavior analytics in real time, using smartphone device sensors. Safety efficiency is examined in terms of speed limit violation, driving distraction, aggressiveness and safety on urban, rural and highway road and in an overall model. DEA models are identifying the most efficient drivers that lie on the efficiency frontier and act as peers for the rest of the non-efficient drivers. The proposed methodological framework is tested on data from fifty-six (56) drivers during a 7-months period. Findings help distinguish the most efficient drivers from those that are less efficient. Moreover, the efficient level of inputs and outputs to switch from non-efficiency to the efficiency frontier is identified. Results also provide a potential for classification of the driving sample based on drivers' comparative safety efficiency. The main characteristics of the most and less efficient drivers are consequently analyzed and presented. Most common inefficient driving practices are identified (aggressive, risky driving, etc.) and driving behavior is comparatively evaluated and analyzed.

1. Introduction

Measuring driving efficiency has been the focus of many studies in driving behavior literature in the past (Matthews et al., 1996, 1998; Young et al., 2011). From a traffic safety perspective, it is a matter of great significance to identify the parameters that influence driving behavior and therefore traffic risk. Several studies have been carried out regarding mobile phone usage distraction and methodologies for collecting and analyzing (Tselentis et al., 2017) driving behavior data. The most common methodology applied included driving simulators (Desmond et al., 1998; Lenné et al., 1997), questionnaires (Matthews et al., 1998) combined with simulators and naturalistic driving experiments (Toledo et al., 2008; Birrell et al., 2014), while the most common method of monitoring driving measures included recorders that relate to the car engine (Zaldivar et al., 2011; Backer-Grøndahl & Sagberg, 2011) and smartphones (Vlahogianni and Barmpounakis, 2017). As shown from previous research (Eftekhari and Ghatee, 2016; Kanarachos et al., 2018; Gadziński, 2018; Bejani, & Ghatee, 2018; Huang et al., 2019), smartphones and their sensors are increasingly used as devices for monitoring driver behaviour because they present many advantages due to high market penetration rates and Internet of Things (IOT) connectivity.

* Corresponding author.
E-mail address: dtsel@central.ntua.gr (D.I. Tselentis).

https://doi.org/10.1016/j.trc.2019.11.006
Received 11 June 2018; Received in revised form 9 November 2019; Accepted 10 November 2019
0968-090X/

1.1. Human factors in road safety

Regarding mobile phone usage while driving, literature has shown that it has a significant effect on driver behavior. Cell phone use causes drivers to have higher variation in accelerator pedal position, drive more slowly with more variation in speed and report a higher level of workload (Haque and Washington, 2015) regardless of conversation difficulty level. Drivers tend to select larger vehicle spacing (Nilsson, 1982), and longer time headways (Saifuzzaman et al., 2015) suggesting possible risk compensatory behavior (Haque and Washington, 2015; Törnros and Bolling 2006). Furthermore, the participants' reaction times (Patten and Kircher, 2004) increase significantly when conversing, but no benefit of hands-free units over handheld units on rural roads/motorways were found (Handel et al., 2014; Yannis et al., 2014).

Speeding is also recognized as one of the most important factors in driving risk since it influences the accident probability (e.g. decreased reaction distance, loss of control) as well as the crash impact (Mesken et al., 2002). According to (OECD, 2006) speeding has been a contributory factor in 10% of the total accidents and more than 30% in fatal accidents. According to Andersson and Nilsson (1997), Nilsson (1982) the probability of a crash involving an injury is proportional to the square of the speed, the probability of a serious crash is proportional to the cube of the speed and the probability of a fatal crash is related to the fourth power of the speed. Moreover, (Nilsson 2004) depicts the relationship between speed and driving risk via an exponential curve, showing that the driving risk is not proportional to the speed.

Harsh acceleration (HA), harsh braking (HB) and harsh cornering (HC) events are three significant indicators for driving risk assessment (Tselentis et al., 2017; Johnson and Trivedi 2011; Bonsall et al., 2005) especially for evaluating driving aggressiveness. This is because they are strongly correlated with unsafe distance from adjacent vehicles, possible near misses, lack of concentration, increased reaction time, poor driving judgement or low level of experience and involvement in situations of high risk (e.g. marginal takeovers). The correlation between HA and HB events with driving risk has been highlighted in the scientific papers published by (Tselentis et al., 2017; Bonsall et al., 2005) and it has been widely recognized by the insurance and telematics industry (Tselentis et al., 2017).

1.2. Data envelopment analysis (DEA)

The terms “efficiency” and “productivity” are widely used in economics and refer to the optimal way a production unit can make use of its available resources (Shone, 1981). More specifically (Farrell, 1957), a Decision-Making Unit (DMU) is “technically efficient” when the amount of outputs produced is maximized for a given amount of inputs, or for a given output the amount of inputs used is minimized. Thus, when a DMU is technically efficient, it operates on its production frontier and therefore DMUs lie on the efficiency frontier (Ramanathan, 2003). Based on the assumptions that will be stated below, drivers are considered those DMUs whose efficiency is evaluated in this study and DEA applicability on the field of driver’s assessment based on microscopic behavioral characteristics is investigated.

Efficiency can be defined as the ratio of input and output in a theoretical scenario of units that have a single input and output, but in a real case scenario where typical organizational units have multiple and incommensurate inputs and outputs a more scientific approach is needed. Data Envelopment Analysis (DEA) is an approach for efficiency and productivity analysis of production units with multiple inputs to produce multiple outputs mostly used thus far in business, economics, management and health. The rationale for using DEA is its applicability to the multiple input–output nature of DMUs provision and the simplicity of the assumptions underlying the method. It is a methodology of several different interactive approaches and models used for the assessment of the relative efficiency of DMU and for the assessment of the efficiency frontier. It assists in drawing important conclusions on operational management of the efficient and inefficient units.

DEA is a technique of mathematical programming problem with minimal assumptions that determines a unit’s efficiency based on its inputs and outputs and compares it to other units involved in the analysis (Ramanathan, 2003). It is a data-oriented methodology that effects performance evaluations and other conclusions drawn from the analysis directly from the observed data. The efficiency of a DMU is comparatively measured and analyzed relatively to the rest of the DMUs considering that all DMUs lay on or below the efficiency frontier. No assumption is required about functional form (e.g. a regression equation, a production function, etc.) or the statistical distribution of data sample and as a result DEA is classified as a non-parametric method (Ramanathan, 2003). It is a frontier analysis, a process of extremities, not driven by central tendencies in contrast to all statistical procedures. Each DMU is analyzed separately and the real and optimal performance that can be achieved for each unit is estimated.

DEA has become one of the most popular fields in operations research, with applications involving a wide range of context (Thanassoulis, 2001). It has been applied in great extent in literature (Cook & Seiford, 2009; Emrouznejad et al., 2008; Hollingsworth et al., 1999) to measure and compare the productivity performance of a group of DMUs. It is one of the most popular fields in operations research (Emrouznejad et al., 2008; Seiford, 1997) to say the least. Martić et al. (2009) presented the ample possibilities for using DEA for evaluating among others the performance of banks, schools, university departments, farming estates, hospitals and social institutions, military services and entire economic systems. Since the introduction of CCR model (Charnes et al., 1978) in 1978, the number of publications where DEA is implemented has exponentially grown. DEA has also been implemented in transport fields in assessing public transportation system performance (Karlaftis et al., 2013), as well as traffic safety studies (Egilmez & McAvoy, 2013; Alper et al., 2015) where it was proved to be equally useful as in the fields stated above.

DEA is a non-parametric approach that does not require any assumptions about the functional form of a production function and a priori information on importance of inputs and outputs. DEA allows each DMU to choose the weights of inputs and outputs which maximize its efficiency. The DMUs that achieve efficiency equal to unit are considered efficient while the other DMUs with efficiency scores between zero and unit are considered as inefficient. The first DEA model proposed by Charnes et al. (1978) is the CCR model that assumes that production exhibits constant-returns-to-scale (CRS) i.e. outputs are increased proportionally to inputs. DEA models can also be distinguished based on the objective of a model; that can be either outputs maximization (output-oriented model) or inputs minimization (input-oriented model).

1.3. Scope of this study

The concept of DEA (Ramanathan, 2003) is to minimize inputs (input-oriented model) or maximize the outputs of a problem (output-oriented model). More specifically in the case study examined herein, a driver should either drive more kilometers maintaining the same number of harsh braking/accelerating events or reduce the number of harsh braking/accelerating events for the same mileage. The same applies, of course, to the rest of the metrics recorded for each driver. From a road safety perspective, increasing mileage increases the exposure of a driver and consequently crash risk (Tselentis et al., 2017) and, therefore, an input-oriented (IO) DEA model is developed aiming to minimize inputs (recorded metrics) maintaining the same number of outputs (recorded distance). It is also assumed for the sake of simplicity that the driving efficiency problem is a CRS problem (Ramanathan, 2003) and that the sum of all metrics (inputs) recorded such as the number of harsh acceleration and braking events occurred in each trip_{i} changes proportionally to the sum of driving distance (output).

Although a driver is not a decision-making unit with the same sense as the term appears in business and economics research (Shone, 1981), it can be evaluated as such and therefore it will be considered as a DMU for the purpose of this research. This is deemed to be a correct assumption on a driver basis since (a) all variables used are continuous quantitative variables as those used in previous DEA studies (Cook & Seiford, 2009; Hollingsworth et al., 1999; Karlaftis et al., 2013; Egilmez & McAvoy, 2013) and (b) a driver should reduce his mileage (Tselentis et al., 2017) and the frequency of some of his driving characteristics such as harsh acceleration and braking, mobile phone usage and speeding (Tselentis et al., 2017; Aarts & Van Schagen, 2006; Young & Regan 2007).

Driver's safety analysis on a microscopic level has been studied in a great extent in the past (Karlaftis et al., 2013; Orfila et al., 2015; Eboli et al., 2016; Mantouka et al., 2018; Jia et al., 2019; Papadimitriou et al., 2019; Sun et al., 2019). It has also been studied by making use of Data Envelopment Analysis (DEA) techniques applied on simulator data (Babaee et al., 2014a, 2014b, 2015, 2016) but never thus far by applying these techniques on naturalistic driving data. This paper proposes a methodological framework a) for measuring driver's efficiency in terms of safety and categorize the drivers of the sample used in three groups i.e. non-efficient, weakly efficient, most efficient, which are the most common groups developed when applying DEA (Cooper et al., 2006) and b) for estimating the efficient level of driving metrics that each driver should reach to become efficient. The main characteristics of each group are presented in order to draw important conclusions on the features of each driving group and provide recommendations for drivers on how to improve their driving efficiency. In this study, drivers are the Decision Making Units (DMUs) (Ramanathan, 2003) that make decisions for a given mileage range about the number of events occurring, the time of mobile phone usage and speed limit violation and to whom a relative driving safety efficiency index is assigned. Driving attributes (metrics and distance recorded) will be considered the inputs and outputs of the DEA application, which will quantify relative driving performance through estimating driving efficiency. More details on the DEA structure implemented herein are given below. The proposed methodology is applied to a case study of 34,060 recorded trips from fifty-six (56) drivers collected from a naturalistic experiment. For brevity reasons, from now on driving safety efficiency will be mentioned as driving efficiency everywhere in text.

2. DEA and driving efficiency problem

2.1. Mathematical formulation of DEA for the driving efficiency problem

For the sake of simplicity, it is noted that from now on DMUs will be referred as drivers. In order to evaluate the driving efficiency of Driver_0 and assuming a sample of N drivers, let X and Y represent the set of inputs and outputs respectively, for the rest of the drivers' sample. In other words, $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i\}$ and $Y = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_i\}$ where $i \in [1, N - 1]$. The input-oriented CCR model evaluates the efficiency of Driver_0 by solving the linear program (Ramanathan, 2003) presented below. Considering each driver as a DMU and taking into account the principles of DEA (Charnes et al., 1978), the mathematical formulation for the specific driving efficiency problem examined herein is:

$$\min(Driving\_Efficiency_0)$$

Subject to the following constraints:

$$\begin{aligned} &Driving\_Efficiency_0 * x_0 - X * \lambda \geqslant 0 \\ &Y * \lambda \geqslant y_0 \\ &\lambda_i \geqslant 0 \ \forall \ \lambda_i \in \lambda \end{aligned} \tag{1}$$

where λ_i is the weight coefficient for each Driver_{i} that is an element of set λ, X is the set of Inputs (number of harsh acceleration/braking events, seconds of mobile phone usage, seconds driving over the speed limits), Y is the set of Outputs (distance travelled) and $Driving\_Efficiency_0$ is a scalar representing the efficiency of reference driver (the one with Driver ID = 0) i.e. Driver_0. Apparently, the use of the sets in the constraints indicates the creation of (N - 1) inequalities when "building" the constraints of the linear problem i.e. $1 + 3 * (N - 1) = 3 * N - 2$ constraints. The rationality behind these constraints is to ensure that, compared to the rest of the sample, there could not be any other X, Y combination leading to a higher efficiency than that of the driver being evaluated. The set of λ estimated from the linear program is positive only for those drivers who act as peers to the driver being evaluated and is used afterwards to estimate the efficient level of inputs for the inefficient drivers (driving efficiency < 1) that each driver should reach to become efficient. The objective function of DEA is min $\mathrm{Driving\_Efficiency}_i$ i.e. to determine the minimum efficiency of driver_i that satisfies the above conditions. This formulation denotes that the target of this linear problem is to find the minimum driving efficiency that satisfies all the above constraints and not to minimize driving efficiency itself. To benchmark the efficiency of each and every driver in the database, this linear programming (LP) problem should be solved for each driver_i of the sample or in other words, N times in total.

2.2. Efficient level of inputs and outputs for non-efficient drivers

After DEA LPs of (1) are solved and the efficiency index $Driving\_Efficiency_B$ and coefficients λ_i are estimated for each driver the efficient level of inputs and outputs at which each driver could optimally reach can be calculated. The efficient level of inputs for driver i can be calculated as the product sum of the lamdas and the input values of each of the identified peers whereas to find the efficient level of outputs for the same driver, each output value should be divided by theta value. Considering $driver_i$ as the reference DMU and a set of m drivers, where $m \in \mathbb{N}$ is the number of $driver_i's$ peers, the efficient level of $Metric_i$ can be estimated using following formula (2):

$$Metric_i = \sum_{j=1}^{m} \lambda_j * Metric_j \tag{2}$$

More specifically, considering $driver_i$ as the reference DMU and a set of m drivers, where $m \in \mathbb{N}$ is the number of $driver_i's$ peers, the efficient level of e.g. ha_{urban} can be estimated using following formula (3):

$$ha_{urban_i} = \sum_{j=1}^{m} \lambda_j * ha_{urban_j} \tag{3}$$

On the other hand, the efficient level of e.g. $distance_{urban}$ is calculated from formula (4):

$$distance_{urban} = distance_i / Driving\_Efficiency_i \tag{4}$$

It should be noted that a DMU achieves its efficient level by reaching the efficient level of either its inputs or outputs. Additionally, a DMU is deemed to have achieved the efficient level when it reaches unit efficiency (Ramanathan, 2003). Based on the above, it can be concluded that the required change of each driving attribute that was taken into consideration in order for a driver to shift either to the efficient frontier or to another driving class (group of drivers with different driving average safety efficiency) can be estimated. This can be achieved by solving the optimization problem for a specific input or output given the target efficiency ($Driving\_Efficiency_B$), which is the upper or the lower limit of the class that the driver is shifting in case of efficiency decrease or increase respectively.

3. Experimental data collection

The implementation of the proposed methodology is based on the use of indicators to describe risk exposure driving performance extracted from the driving trips data set of OSEVEN insurance telematics and driving behavioral analytics platform (www.oseven.io). The main risk exposure indicators are:

- total distance travelled (between the start point and the end point of the trip),
- driving duration (time difference between the trip start time and trip end time),
- type(s) of the road network used (urban, rural, highway), and
- time of the day driving (morning peak/rest of the day).

It is clarified that at the present study, those motorways with a speed limit of 90 km/h and higher are taken into account as highways. The above are combined with other data sources (e.g. google maps service). The main driving behavior indicators are:

- speeding (percentage of time driving over the speed limits),
- mobile phone use while driving (identifying any movement of the smartphone device),
- number and severity of harsh events:
  - o harsh braking,
  - o harsh acceleration and
  - o harsh cornering.

The above indicators are estimated by sensing and fusing Accelerometer*, the Gyroscope*, Magnetometer and the GPS (speed, course, longitude, latitude) data from smartphones in a nonintrusive manner using the OSEVEN proprietary algorithms. The recording is done using a dedicated smartphone app in 1 Hz. The application developed detects when someone is driving and

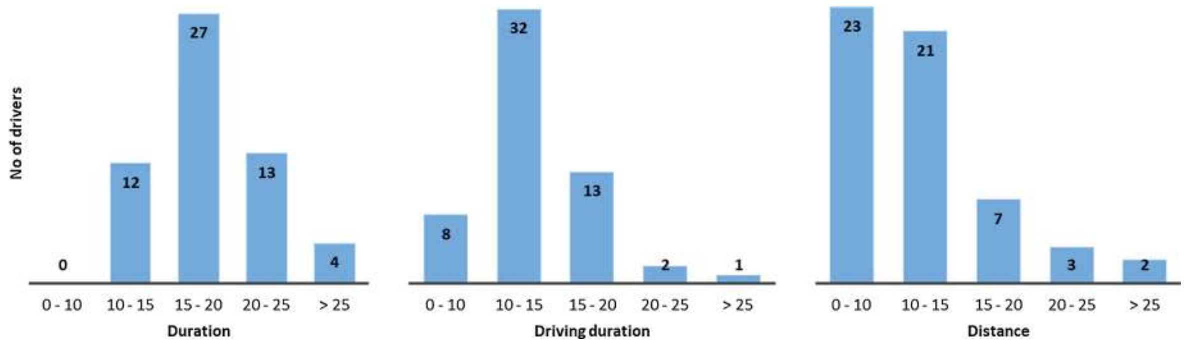

Fig. 1. Histogram of the (i) average trip duration (min), (ii) average trip driving duration (min) and (iii) average trip distance (km) of the driving sample (from left to right).

automatically initiates the recording procedure. All data are anonymized before obtained from OSeven so that there is no possibility to create a connection between subjects and driving data since data related to (a) personal information (name, age, gender, email address, etc.) or (b) geographical information data are not provided. All terms of use are described in the OSeven application when it is initially downloaded. It should also be highlighted at this point that the approach followed in this study aims to identify driving behaviors and the factors influencing them and not explain the causality between behavior and other factors such as age, gender, occupation etc. or describe the distribution of the driving sample collected. The advantage of such an approach is that behaviors can be studied even in cases where demographic data of a driving sample are not available or cannot be collected.

For the purposes of this research, driving data from two hundred and thirty six (236) drivers were randomly selected from OSeven database, constituting a large database of 50,741 trips. For each driver, all trips that took place between August 2016 and April 2017 were selected. The first criterion chosen by the authors for specifying the driver's sample were adopted from study (Shichrur et al., 2014) which proved that sampling less than 100 driving hours per driver does not result in a reliable measure for analyzing driving patterns and changes in the behavior of drivers over time. On the top of that, all drivers should have positive mileage on all three types of road network. The third criterion was that drivers with zero input attributes (i.e. zero harsh acceleration, braking, speed limit violation, mobile phone usage) should be eliminated from the sample which is a limitation of DEA. The business equivalent of a zero input could be a factory that is producing a product without making use of any material and/or workforce which practically cannot occur. For the same reason, harsh cornering events were finally eliminated as a DEA input, as, in most cases, there are no such events in highways and, therefore, results would not be comparable. This procedure resulted to 56 drivers who fulfilled these 3 criteria and were further analyzed. The total number of trips that were conducted by the 56 drivers was 34,060 constructing thus a large database.

Figs. 1 and 2 and Table 1 illustrate some descriptive statistics regarding the attributes of the driving sample collected from the smartphone devices. The first figure presents the average duration, average driving duration (duration of a trip with no stops included) and average distance travelled respectively while the second presents the average number of ha events occurred in urban, rural and highway road network per 100 km distance travelled in urban, rural and highway network respectively. Table 1 provides some descriptive statistics of the cumulative per driver values of the variables recorded. In other words, the final database includes the cumulative value (for the period each driver was recorded) of each variable considered in the DEA models, constructing thus a database with one row per driver, the descriptive statistics of which are presented in Table 1.

Data processing and DEA improvement algorithms (Tselentis et al., 2019) are performed in Python programming language and several scripts are written for this reason. Python packages used include pandas and numpy for numeric calculations and transformations, scipy that features quickhull algorithm and pulp for LP problem construction. More details on the algorithm implementation

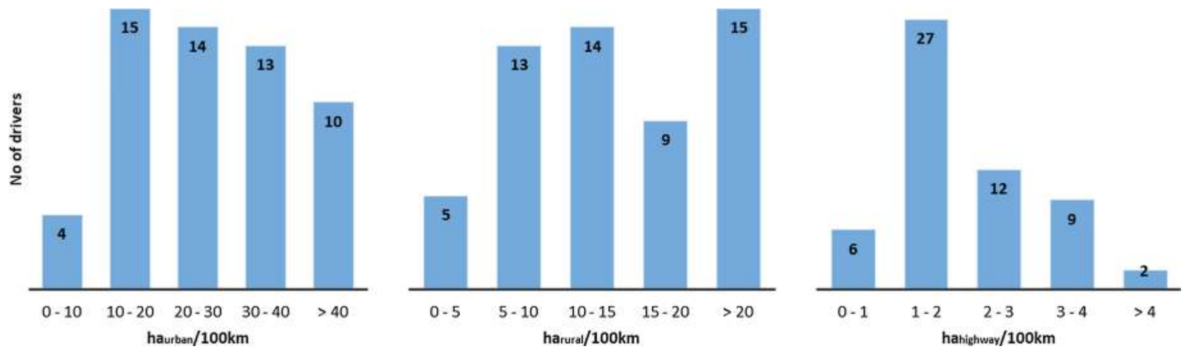

Fig. 2. Histogram of the (i) average ha_{urban}/$distance_{urban}$, (ii) average ha_{rural}/$distance_{rural}$, (iii) average $ha_{highway}$/$distance_{highway}$ per 100 km of the driving sample (from left to right).

Table 1
Descriptive statistics of the cumulative per driver values of the variables recorded.

	Distance (km)	HA	HB	Mobile (secs)	Speeding (secs)
	Urban				
Min	288	31	18	902	3971
Max	5224	2033	1033	84,677	81,640
Average	2238.6	633.8	226.7	18109.7	26134.3
Standard Deviation	1061.8	485.4	184.5	17547.4	15164.7
Median	2166.5	477.5	150.5	11,403	23621.5
Kurtosis	0.5	1	5.3	3.7	1.9
Skewness	0.7	1.2	1.9	1.8	1.1
	Rural				
Min	391	45	14	39	1802
Max	7209	1566	1213	65,623	68,405
Average	2274.6	342.5	173	10735.7	21214.3
Standard Deviation	1363.4	308.4	198.3	14155.8	14862.7
Median	2165.5	235	108	6420.5	19,547
Kurtosis	3.2	3.8	15.5	7.5	0.7
Skewness	1.5	1.8	3.6	2.8	1
	Highway				
Min	180	2	1	37	23
Max	6932	185	98	27,979	48,367
Average	1512.7	30.3	16.4	2841.8	9126.4
Standard Deviation	1209.5	35.2	18	4384.8	10204.4
Median	1147.5	16	9.5	1604.5	5357.5
Kurtosis	6	8.4	7.4	19	4.1
Skewness	2	2.8	2.4	3.8	1.9

are given below. Coding is applied using Pycharm IDE Community edition, for Python & Scientific development. The computer used is an Intel® Core™ i7 CPU K 875 @ 2.93 GHz × 8 featuring a 2.0 GiB Ram memory running on ubuntu 16.04 LTS.

4. Implementation and results

4.1. Input and output selection

Models representing driving behavior in all road types and in total are developed with multiple inputs and outputs. A critical process for DEA is input and output selection. Thus, selection process should be linked to the conceptual specifications of each problem. Several issues that should be taken into consideration before applying DEA to a dataset are discussed in Dyson et al. (2001). One of the pitfalls is that the efficiency score might be miscalculated when input and output variables are in the form of percentiles and/or ratios simultaneously with raw data (Cooper et al., 2006). Taking this into account the specific data used in this study are metrics recorded in the form of raw data i.e. the number of harsh braking, accelerations and cornering events, seconds driving over the speed limit and seconds used the mobile phone and not as ratios or percentiles e.g. percentage of distance driving over the speed limit. In the particular DEA formulation of this research, the driving metrics (number of harsh acceleration/braking events, mobile phone usage, speed limit exceedance) are used as inputs and the distance is used as output. Therefore, this method would have taken into account twice the parameter of distance if the normalized measures of the inputs over the distance units were used. This methodology assesses comparatively the driving safety efficiency taking into account the metrics recorded (inputs) during the specific distance recorded (output). As shown in the previous sections of this paper, literature review revealed that all these indicators are the most influencing factors of accident risk that is the reason why they are used in the models implemented. The main reason for performing the analysis in each road type separately, where driving conditions exhibit homogeneous characteristics, is because of the different road and traffic characteristics that each type presents. All indicators along with distance travelled by drivers are recorded per road type (urban, rural, highway) and in total e.g. number of harsh accelerations that occurred in urban road, time violating speed limits etc. Variables used in the analysis along with their description are shown in Table 2.

The driver is deemed to be a DMU with an aggregate performance for the entire monitoring period. Moreover, his driving behavior is considered equivalent to the sum of the driving characteristics that were recorded for the period examined. For instance, the total distance travelled in urban network is equivalent to the sum of the distance travelled in urban network in each $trip_{ij}$ (where i is the index of $driver_i$ and j the index of $trip_j$ of $driver_i$) by the specific $driver_i$ (DMU_i). In general, the same applies for all indicators of $driver_i$, which are calculated aggregately as shown in the following formula (5):

$$indicator_i = \sum_{j=1}^{N_i} indicator_{ij} \tag{5}$$

recorded $\forall\ trip_j,\ j \in (1, N_i)$ that took place by $driver_i$. As described above, each driver is treated as a distinct DMU to be analyzed in DEA and therefore the linear program constructed (see (1)) has 57 variables (λ_i, θ_B) that is equal the number of drivers plus the efficiency for $driver_0$. The number of constraints on the other hand is equal to the sum of a) the number of inputs

Table 2
Description of the per trip variables recorded.

Variable name	Variable short description
ha_X	Number of harsh acceleration events in X road type
ha_{urban}	Number of harsh acceleration events in urban road
ha_{rural}	Number of harsh acceleration events in rural road
$ha_{highway}$	Number of harsh acceleration events in highway
hb_X	Number of harsh braking events in X road type
hb_{urban}	Number of harsh braking events in urban road
hb_{rural}	Number of harsh braking events in rural road
$hb_{highway}$	Number of harsh braking events in highway
$speeding_X$	Total seconds of speed limit violation in X road type
$speeding_{urban}$	Total seconds of speed limit violation in urban road
$speeding_{rural}$	Total seconds of speed limit violation in rural road
$speeding_{highway}$	Total seconds of speed limit violation in highway
$mobile_X$	Total seconds of mobile phone usage in X road type
$mobile_{urban}$	Total seconds of mobile phone usages in urban road
$mobile_{rural}$	Total seconds of mobile phone usage in rural road
$mobile_{highway}$	Total seconds of mobile phone usage in highway
$distance_X$	Total distance driven in X road type
$distance_{urban}$	Total distance driven in urban road
$distance_{rural}$	Total distance driven in rural road
$distance_{highway}$	Total distance driven in highway

($\theta_B * \chi_0 - X * \lambda \geqslant 0$), b) the number of outputs ($Y * \lambda \geqslant y_0$) and c) the number of drivers ($\lambda_i \geqslant 0$).

The DEA procedure described by (1) is followed separately for each of the three different road types (urban, rural, highway) and aggregately in an overall model as described in Table 3. Data filtering and DEA improvement algorithms are performed in Python programming language and several scripts are written for this reason. Python packages used include Pandas and Numpy for numeric calculations and transformations and Pulp for LP problem construction.

This results to 16 different models of which 12 are per road type and 4 overall. The variables' combinations for structuring the four models of each category was based on the literature review conducted above. Model 1 and 2 represents the speed limits violation and mobile phone distraction. Model 3 incorporates the three most significant explanatory driving indicators for driving aggressiveness, while model 4 is the overall model that includes all traffic safety parameters found in literature review and accounts for the overall safety profile of the driver.

Fig. 3 illustrates the results of model 3 per road type where as it appears there is only one efficient driver for urban and rural road, whereas for highway there are two, which confirms the results of the DEA LPs. In every subplot of Fig. 3, $distance_x/ha_x$ and $distance_x/hb_x$ is plotted in axis Y and X respectively along with the envelopment line accounting for the efficiency frontier. Extending the line joining the origin and DMU_i, it crosses the efficiency frontier at a point where virtual DMU_i' is created which represents the optimal performance which the specific DMU_i can achieve. The closer a driver is to the efficiency frontier, the higher its efficiency index is. In urban and rural road subplots, the influence of outliers to the DEA solution is obvious since most drivers appear to be near the origin. Nonetheless, the solution remains reliable as the efficiency index calculated is comparable to that of the rest of the drivers'

Table 3
Inputs and Outputs of the DEA models used.

	Per road type model		Overall model		
	Set of Inputs used	Set of Outputs used	Set of Inputs used		Set of Outputs used
Model type 1	(1) $speeding_x$	(1) $distance_x$	(1) $speeding_{urban}$ (2) $speeding_{rural}$ (3) $speeding_{highway}$		(1) $distance_{urban}$ (2) $distance_{rural}$ (3) $distance_{highway}$
Model type 2	(1) $mobile_x$	(1) $distance_x$	(1) $mobile_{urban}$ (2) $mobile_{rural}$ (3) $mobile_{highway}$		(1) $distance_{urban}$ (2) $distance_{rural}$ (3) $distance_{highway}$
Model type 3	(1) ha_x (2) hb_x	(1) $distance_x$	(1) ha_{urban} (2) ha_{rural} (3) $ha_{highway}$	(4) hb_{urban} (5) hb_{rural} (6) $hb_{highway}$	(1) $distance_{urban}$ (2) $distance_{rural}$ (3) $distance_{highway}$
Model type 4	(1) ha_x (2) hb_x (3) $speeding_x$ (4) $mobile_x$	(1) $distance_x$	(1) ha_{urban} (2) ha_{rural} (3) $ha_{highway}$ (4) hb_{urban} (5) hb_{rural} (6) $hb_{highway}$	(10) $speeding_{urban}$ (11) $speeding_{rural}$ (12) $speeding_{highway}$ (13) $mobile_{urban}$ (14) $mobile_{rural}$ (15) $mobile_{highway}$	(1) $distance_{urban}$ (2) $distance_{rural}$ (3) $distance_{highway}$

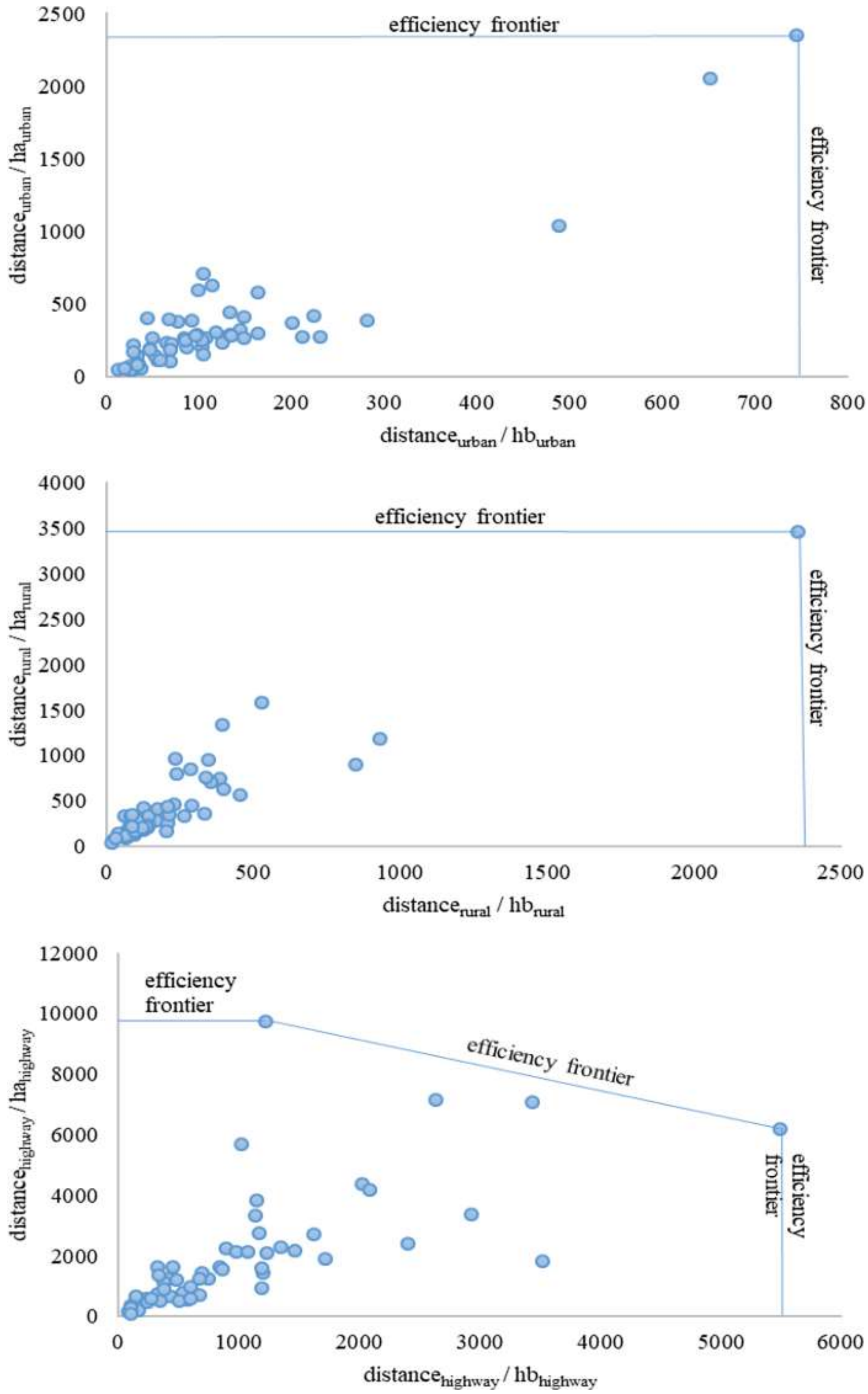

Fig. 3. Efficiency frontier of drivers' aggressiveness for (i) urban, (ii) rural, (iii) highway road types (model type 3) (from left to right).

set. It should be highlighted that models incorporating two-inputs/one output or one-input/two outputs can only be visualized in 2-D figures.

4.2. Drivers sample classification

The results of DEA are the efficiency index *Driving_Efficiency*$_B$ and coefficients λ_i for each driver. This allows for the classification

Table 4
Driving characteristics of efficiency classes per road type and in the entire driving data (overall).

Model		Efficiency classes		
		0–25% percentile	25–75% percentile	75–100% percentile
Urban	1	$speeding_{urban}$ = 20.08%	$speeding_{urban}$ = 11.95%	$speeding_{urban}$ = 6.51%
	2	$mobile_{urban}$ = 19.48%	$mobile_{urban}$ = 6.80%	$mobile_{urban}$ = 2.31%
	3	ha_{urban}/100 km = 45.97	ha_{urban}/100 km = 27.40	ha_{urban}/100 km = 10.71
		hb_{urban}/100 km = 17.38	hb_{urban}/100 km = 8.99	hb_{urban}/100 km = 5.08
	4	ha_{urban}/100 km = 41.06	ha_{urban}/100 km = 22.85	ha_{urban}/100 km = 24.72
		hb_{urban}/100 km = 16.75	hb_{urban}/100 km = 8.43	hb_{urban}/100 km = 6.81
		$mobile_{urban}$ = 17.77%	$mobile_{urban}$ = 6.78%	$mobile_{urban}$ = 4.05%
		$speeding_{urban}$ = 15.79%	$speeding_{urban}$ = 13.02%	$speeding_{urban}$ = 8.66%
Rural	1	$speeding_{rural}$ = 23.79%	$speeding_{rural}$ = 14.21%	$speeding_{rural}$ = 6.33%
	2	$mobile_{rural}$ = 15.10%	$mobile_{rural}$ = 5.69%	$mobile_{rural}$ = 1.64%
	3	ha_{rural}/100 km = 23.65	ha_{rural}/100 km = 14.28	ha_{rural}/100 km = 6.36
		hb_{rural}/100 km = 11.43	hb_{rural}/100 km = 6.96	hb_{rural}/100 km = 3.00
	4	ha_{rural}/100 km = 20.31	ha_{rural}/100 km = 12.32	ha_{rural}/100 km = 13.62
		hb_{rural}/100 km = 8.71	hb_{rural}/100 km = 6.26	hb_{rural}/100 km = 7.13
		$mobile_{rural}$ = 10.28%	$mobile_{rural}$ = 6.51%	$mobile_{rural}$ = 4.81%
		$speeding_{rural}$ = 20.58%	$speeding_{rural}$ = 14.49%	$speeding_{rural}$ = 8.97%
Highway	1	$speeding_{highway}$ = 32.39%	$speeding_{highway}$ = 13.06%	$speeding_{highway}$ = 3.98%
	2	$mobile_{highway}$ = 12.34%	$mobile_{highway}$ = 3.73%	$mobile_{highway}$ = 0.74%
	3	$ha_{highway}$/100 km = 3.40	$ha_{highway}$/100 km = 1.74	$ha_{highway}$/100 km = 0.98
		$hb_{highway}$/100 km = 1.67	$hb_{highway}$/100 km = 1.02	$hb_{highway}$/100 km = 0.49
	4	$ha_{highway}$/100 km = 2.80	$ha_{highway}$/100 km = 1.91	$ha_{highway}$/100 km = 1.24
		$hb_{highway}$/100 km = 1.61	$hb_{highway}$/100 km = 1.05	$hb_{highway}$/100 km = 0.50
		$mobile_{highway}$ = 5.40%	$mobile_{highway}$ = 5.61%	$mobile_{highway}$ = 3.92%
		$speeding_{highway}$ = 29.31%	$speeding_{highway}$ = 13.08%	$speeding_{highway}$ = 7.01%
Overall	1	$speeding_{urban}$ = 17.12%	$speeding_{urban}$ = 12.50%	$speeding_{urban}$ = 8.37%
		$speeding_{rural}$ = 21.25%	$speeding_{rural}$ = 14.41%	$speeding_{rural}$ = 8.48%
		$speeding_{highway}$ = 24.24%	$speeding_{highway}$ = 14.26%	$speeding_{highway}$ = 9.72%
	2	$mobile_{urban}$ = 17.07%	$mobile_{urban}$ = 7.22%	$mobile_{urban}$ = 3.89%
		$mobile_{rural}$ = 13.30%	$mobile_{rural}$ = 5.99%	$mobile_{rural}$ = 2.85%
		$mobile_{highway}$ = 9.75%	$mobile_{highway}$ = 4.37%	$mobile_{highway}$ = 2.05%
	3	ha_{urban}/100 km = 36.94	ha_{urban}/100 km = 30.09	ha_{urban}/100 km = 17.13
		ha_{rural}/100 km = 19.26	ha_{rural}/100 km = 16.26	ha_{rural}/100 km = 8.46
		$ha_{highway}$/100 km = 3.12	$ha_{highway}$/100 km = 1.76	$ha_{highway}$/100 km = 1.32
		hb_{urban}/100 km = 12.42	hb_{urban}/100 km = 10.34	hb_{urban}/100 km = 7.87
		hb_{rural}/100 km = 9.33	hb_{rural}/100 km = 7.36	hb_{rural}/100 km = 4.85
		$hb_{highway}$/100 km = 1.44	$hb_{highway}$/100 km = 0.95	$hb_{highway}$/100 km = 0.87
	4	–	–	–

of the whole set of drivers to most efficient, weakly efficient and non-efficient. Since the absolute value of the efficiency index cannot be somehow interpreted unless it is compared to the efficiency index of the rest of the drivers' set, the percentiles of the driver sets $Driving\_Efficiency_B$ are used to classify drivers. The percentile thresholds specified was 25% and 75%, which separate the subsets of non-efficient and weakly efficient as well as weakly efficient and most efficient drivers respectively. The average of the attributes of each class arising, weighted on the distance (for harsh acceleration and braking) or driving time (for speeding and mobile usage) travelled by each driver, are shown in Table 4 where the models per type, road type and overall are presented based on the inputs that were used in each model. For brevity purposes, from here on class 1 drivers will be referred to as most efficient drivers even though only drivers with unit efficiency lie on the efficiency frontier.

For instance, in model $Rural_3$ (representing model 3 of rural road type) the average ha_{rural} and hb_{rural} per 100 km travelled (ha_x, hb_x are the inputs of model 3 for every road type as shown in table 3) of each class are illustrated. For better understanding, results are presented as a percentage of driving time for speeding and mobile usage and as events per 100 km driven for harsh acceleration and braking.

As expected for models 1, 2 and 3 in every road type the average of the attributes is reducing while a driver becomes more efficient. The reason why this is not valid for model 4 of urban and rural road types is probably because (a) while the number of inputs and outputs increases, the number of efficient drivers are increasing as well, especially for small scale samples as the one examined a which renders the classification of the drivers to be more difficult and less accurate since many drivers have unit efficiency and (b) of DEA's sensitivity to outliers, which means that the model can sometimes be influenced by the extreme values of other inputs or outputs e.g. low values of speeding or mobile usage when estimating a driver's efficiency.

Another observation is that the number of harsh events occurring in urban road is extremely higher than in rural and highway and that the number of harsh events in rural road is higher than in highway. The same is noticed for mobile usage but not for speeding where apparently, drivers of all classes tend to drive over the speed limits in rural and highway at least the same or more than in urban. As for model 4 of all road types it should be highlighted that for a specific class some attributes appear to be higher compared to model 1, 2 or 3 probably because more parameters are taken into account in the model that might affect the final configuration of each class.

In general, it can be concluded from model 1 that speed limit violation does not fluctuate and is limited to less than 6.5% of driving time for most efficient drivers in all road types whereas for non-efficient drivers it ranges from 20% to over 32%. As for the set of weakly efficient drivers speed limit exceedance is around 12%–14%. In terms of mobile usage distraction, it appears that non-efficient drivers use their mobile phone significantly more than the other two classes averaging at 16% while most efficient drivers use it less than 1.5% in average. Finally, weakly efficient group of drivers make mobile usage of less than 7%.

It is also noticeable from model 3 that drivers of all ranges of aggressiveness have a 2–3 times larger number of harsh acceleration than braking events per 100 km of driving. For instance, in urban roads, the number of harsh acceleration events ranges from 11 to 46 per 100 km while the number of harsh braking events from 5 to 17.4 for most efficient to non-efficient drivers. The ranges become narrower for rural and highway. In terms of traffic safety, the conclusion that can be drawn from model 4 is that the overall driving profile of a "safer" driver in urban and rural road is not considerably influenced by the driver's number of harsh events since it is much higher than in model 3 where it accounts for aggressiveness. On the other hand, in highway, mobile usage and speeding seems to be significantly higher than model 1 and 2 whereas the number of harsh acceleration and braking events appears to be more critical since they are kept at a much lower level. The same is observed in highways for weakly efficient drivers but not for non-efficient who tend to have a lower mobile usage rate than in model 2, which accounts for distraction. Additionally, weakly efficient drivers in urban and rural road have a lower average number of harsh acceleration event and in average, the same driving characteristics for the rest of the attributes investigated. Finally, for non-efficient drivers of urban and rural road, it was found that all driving attributes were reduced compared to model 1, 2 and 3 probably due to the interaction among variables.

As stated above, as the number of inputs and outputs increases while the number of DMUs remains low, the number of efficient DMUs that are found to be efficient is radically increased. This is the case of the overall model, model 4, where 38 drivers with unit efficiency were found and this is the reason why the authors did not consider it to be significant enough to be presented.

When considering all road types together in Table 4, in terms of speeding percentages a greater tolerance is noticed for drivers to be characterized as most efficient or weakly efficient than in per road type models, which appear to be from slightly in class 2 rural to more than 100% more in class 3 highway model. The same is observed for model 2 and 3 as well for class 2 and 3 drivers except for $hb_{highway}$ which are slightly lower in the overall model. On the other hand, non-efficient drivers have lower speeding percentages in all road types and especially in highway where the difference is higher. The same can be highlighted for model 2 and 3 in highway.

Fig. 4 illustrates the distribution of driving efficiency among the three different road types examined. The distribution of driving efficiency in urban appears to be normally distributed although there are no observations in the first value range. As for rural roads, a concentration of values is observed at the middle and higher values of the graph whereas for highways, this concentration is observed at the lower and higher values.

Efficient level of inputs and outputs for non-efficient drivers

Table 5 shows lambdas and theta for the first twelve drivers, where L_x stands for the lambda coefficient of the efficient driver x that acts as a peer for the DMU examined. For the purpose of brevity, not all lambdas and thetas calculated are presented herein. For instance, for the first row of the table where DEA is solved for $driver_1$, the value of the theta coefficient is 0.581 (less efficient) and lambda coefficients L_{12}, L_{34}, L_{40} and L_{42} of $driver_1$ are equal to 0, 0.52, 0.14 and 0.06 respectively. The efficient level of inputs for $driver_1$ can be calculated as the product sum of the lambdas and the input values of each of the identified peers whereas to

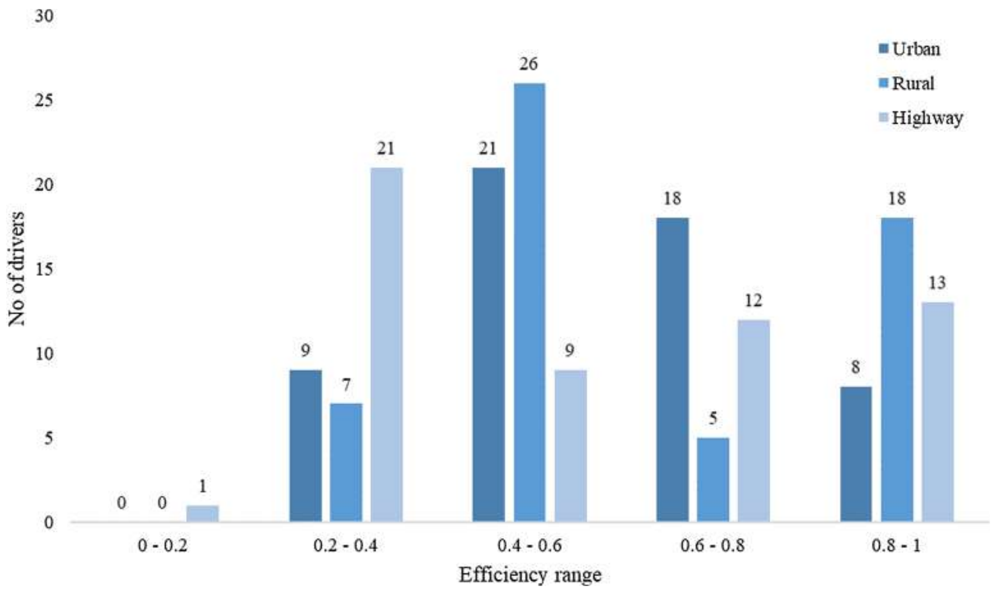

Fig. 4. Histogram showing the frequency of each efficiency range for all 3 road types.

Table 5
Lamdas, thetas, real and efficient level of metrics (distance (km), ha, hb, speeding (sec), mobile (sec)) for the first 12 drivers (DMUs).

	Real level of metrics						Lamdas of peers: Driver No				Efficient level of metrics				
Driver No	$distance_{urban}$	ha_{urban}	hb_{urban}	$speeding_{urban}$	$mobile_{urban}$	Theta	12	34	40	42	$distance_{urban}$	ha_{urban}	hb_{urban}	$speeding_{urban}$	$mobile_{urban}$
1	1868	326	134	21,712	9954	0.581	–	0.52	0.14	0.06	3214.3	159.8	77.9	12617.9	5784.7
2	2456	574	85	27,049	13,974	0.696	–	0.19	0.40	0.20	3526.5	154.8	59.2	18838.2	9732.1
3	1634	709	509	15,888	42,817	0.391	0.79	–	0.20	–	4182.6	277.0	78.1	6206.9	10434.2
4	2219	233	181	27,052	12,421	0.637	–	0.23	0.31	0.18	3481.0	148.5	59.1	17244.6	7917.9
5	4223	1088	309	37,825	29,581	0.613	0.30	1.16	0.47	–	6887.3	417.6	189.5	23192.7	18137.9
6	2773	652	251	25,829	25,887	0.529	0.60	0.51	0.30	–	5245.3	344.7	128.8	13654.8	13685.4
7	2086	265	149	20,880	14,036	0.619	–	0.52	0.28	0.04	3371.9	163.9	79.2	12917.2	8683.2
8	1789	460	323	17,185	5960	0.656	–	0.75	–	0.01	2728.0	184.3	98.1	11269.8	3908.5
9	1630	184	82	30,801	8955	0.528	–	–	0.21	0.22	3085.0	97.2	30.8	14784.6	4731.5
10	808	266	95	9985	6562	0.443	0.11	0.24	0.04	–	1824.6	97.2	42.1	4421.6	2905.8
11	3012	913	152	21,604	43,562	0.585	0.72	–	0.83	–	5149.6	308.3	88.9	12636.1	23107.5
12	1462	329	92	5074	7781	1.000	1.00	–	–	–	1462.0	329.0	92.0	5074.0	7781.0

find the efficient level of outputs for the same driver, each output value should be divided by theta value. Again, taking driver 1 as example, the efficient level of ha_{urban} (most values of ha_{urban} do not appear in Table 5) can be estimated using (2):

$$Efficient\ \ Level\ \ of\ \ ha_{urban_1} = L_{12} * ha_{urban12} + L_{34} * ha_{urban34} + L_{40} * ha_{urban40} + L_{42} * ha_{urban42}$$
$$= 0 * 329 + 0.52 * 242 + 0.14 * 86 + 0.06 * 366 = 159.8$$

On the other hand, the efficient level of e.g. $distance_{urban}$ is calculated (results do no match exactly because theta values have been rounded before presented in Table 5) using (4):

$$Efficient\ \ Level\ \ of\ \ distance_{urban_1} = distance_{urban1}/theta_1 = 1868/0.581 = 3214.3$$

It should be highlighted though, that a DMU should reach either the efficient level of inputs or the efficient level of outputs in order to become efficient and not both at the same time. Of course, if a DMU achieves the efficient level of both inputs and outputs it will become the most efficient DMU and, therefore, it will define a new efficiency frontier and act as a peer for the rest of the DMUs (given that no other DMU will achieve the same). It is also obvious from the table that the most efficient drivers of the sample are drivers 12, 34, 40 and 42 who act as peers for the rest of the driving sample. As expected, most peers do not act as peers for all drivers but most drivers have a portion of the most efficient drivers as their peers. It is also expected that $driver_{12}$, that has unit efficiency, has an equal real and efficient level of metrics for all metrics.

4.3. Ranking validation

Since DEA is not a statistical or econometric method for forecasting future observations of a variable/time-series, no validation technique could be applied for this methodology (e.g. cross-validation, testing different datasets, etc.). Nonetheless, in order to investigate the sensitivity of the estimated efficiencies and the resulting rankings, a sensitivity analysis of the results is carried out by studying the impact of removing DMU's (drivers) from the database. This ultimately provides an indication of the stability of the results arising when the database is partially changed. It is highlighted though that this procedure mainly denotes the representativeness of the sample collected rather than the validity of the methodology used.

In order to test the validity of the rankings, and therefore the relative efficiencies estimated, a random set of 6 DMUs (approximately the 10% of the sample) is deducted from the entire sample and the efficiencies and rankings of the rest of the sample are re-estimated. This procedure is repeated 5 times and the Average and Standard Deviation (StD) of the ranking of each DMU is estimated afterwards. The DMU frequency for 5 different StD ranking position categories and 4 different average ranking position categories (i.e. the average and StD of the ranking position of each DMU that resulted from the above repetitive procedure) are presented below in Fig. 5. It is highlighted that for the overall category of the average DMU ranking position, frequencies are estimated without taking into account the average ranking position whereas for the rest, results are filtered based on ranking position i.e. DMUs with average ranking positions $\in [0, 33]$, $(33, 66]$ and $\in (66, 100]$, respectively.

It becomes apparent that DMUs of higher and lower average ranking positions are less affected by the alterations in the driving sample. The part of the sample that is affected the most by these changes are those DMUs that belong to the middle classes of average ranking positions. Nonetheless, even these DMUs do not seem to be significantly affected, since the two thirds have a StD of less than or equal to 3. The percentage of the overall sample that meets this condition (StD $\leq$ 3) is approximately 82%. The x-axis of Fig. 5 represents the 5 different StD ranking position categories, each different bar color one of the 4 different average ranking position

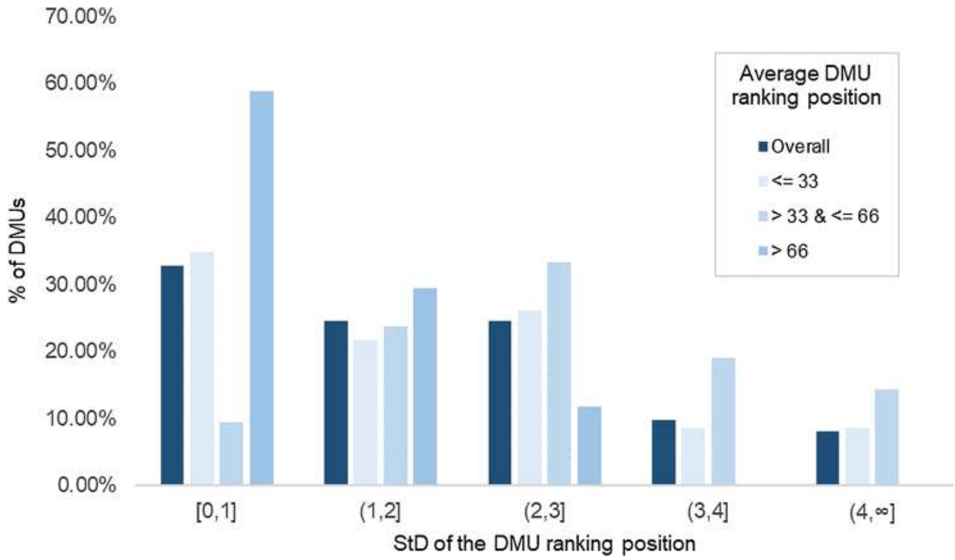

Fig. 5. Histogram showing the DMU frequency for 5 different StD ranking position categories and 4 different average ranking position categories.

categories and the y-axis shows the % DMU frequency that falls into each StD ranking position category grouped by average ranking position category.

5. Discussion

This paper provides an innovative solid framework for benchmarking and evaluation of drivers' efficiency based on Data Envelopment Analysis (DEA). Data exploited were derived from a database of fifty-six (56) drivers, using a sophisticated data collection method from smartphone device sensors, which continuously recorded real time information of driving behavior for 7-months. Combinations of driving analytics collected are taken into consideration for driving assessment, including distance travelled, speed, accelerations, braking and smartphone usage, which serve as inputs and outputs of DEA models that estimate a comparative efficiency index for each driver in the sample. Efficiency is examined in terms of speed limit violation, driving distraction from mobile phone usage, aggressiveness and overall safety on urban, rural and highway roads and in an overall model.

Findings pointed towards a potential for classifying driving sample based on drivers' comparative efficiency is identified. Drivers were divided into three categories (non-efficient, weakly efficient and most efficient) based on the 25% and 75% percentile thresholds specified. The highlights of the analysis conducted for each category indicated considerable differences in driving characteristics between inefficient drivers and the classes of weakly efficient and most efficient drivers with the difference of the two latter to be less significant. Concerning aggressiveness, harsh braking events appeared to be 2–3 times less than harsh acceleration events in all models indicating a higher significance of this attribute for a driver to be characterized as aggressive. The same observation is made for harsh acceleration events in overall safety models (model 4) of all road types where percentage of speeding and mobile usage was identified as key factors for safety efficiency index estimation.

Moreover, the proposed methodology can be used to estimate the optimal level of inputs or outputs that each driver should reach to shift to the efficiency frontier or become even more efficient than those. The latter can potentially serve as a recommendation system's service that provides the appropriate stimuli to drivers to improve their behavior. To this end, gamification policies based on this approach such as competitions, learning goals and awards could also contribute. This could be achieved by a smartphone app that provides feedback to drivers based on their overall or per road type driving efficiency. Drivers could be advised on the driving characteristics that need further improvement for the drivers to become less risky. Moreover, findings could also be useful for developing insurance pricing based on driving usage i.e. Pay-How-You-Drive driving insurance schemes, a policy that also conduces to the further enhancement of behavior and, therefore, driving risk reduction.

The main limitations of this study that are not tackled are summarized below together with several suggestions on how the could potentially be overcome:

First of all, this study takes into account a specific driving time period of a user and not its progression over time, which is equally important when studying driving behavior. Therefore, the temporal dynamics of driving efficiency, should be further investigated and the moving time window in which each driver is assessed is to be specified. It is expected that despite the fact that drivers retain a steady driving behavior for a certain period, there exist dynamic major shifts in systematic behavior within a long-term period. Therefore, drivers should be continuously monitored and reevaluated to capture these shifts and provide personalized advice on how their behavior could be improved in the future. When benchmarking using DEA, the sample should be assessed on a regular basis to identify any possible alterations in the efficiency frontier, which will result in a change in the ranking of the drivers.

The dynamic evolution of driving efficiency raises also the question of how much and how rapidly driving profiles are altering over time. It is a matter of great significance to shed light on this issue and classify drivers based on these characteristics to provide even enhanced recommendations that could potentially reduce driving risk. Another important research question raised at this point is whether future research should focus on the investigation of the macroscopic or microscopic behaviour of drivers. Although these two paths are seemingly different, they are likely to be equally useful in determining the variety of driving behavior patterns. The macroscopic approach would suggest constructing all possible driving profiles and study behavioral shifts among them over longer time periods, whereas microscopic analysis would suggest the opposite i.e. to focus on how everyday driving behavior could be classified as risky or less risky. In any case, future research should focus on the comparison of the results arising from per trip and per driver analysis of each driver to evaluate the representativeness of the results.

Because of DEA's drawback, in terms of the significant required computation time to run, a (hybrid) methodology should be developed to tackle this important issue of running DEA in a timely manner. Additionally, the larger a sample is, the more representative it becomes. Consequently, future research should center to larger samples of trips by collecting a sample of drivers that is representative. The recent trend in driving data collection and analysis is to collect anonymized data from larger samples in contrast with the classic studies, which used to design driving experiments that collect data from a sample the personal details of which, such as demographics etc. are known. Both approaches have several drawbacks and benefits e.g. the fact that the results of a research that has exploited data that cannot be connected with any personal information cannot be generalized in the population. It is therefore important to somehow bridge the gap between these two approaches and retain the advantages of both. In the authors' opinion, this could potentially be achieved by many means such as obtaining larger samples to respect representativeness, collect data from many countries of which drivers have different driving characteristics etc. This should be the objective of future research as well.

It is a fact that models become more representative of the average characteristics of each class as more trips and drivers are aggregated. As the sample grows bigger, it is expected that the high proportion of efficient drivers to the total number drivers will be reduced and that the relative efficiency and the ranking of most drivers will not be significantly affected. Other DEA's limitations should also be addressed which among others include DEA's sensitivity to outliers and that drivers with zero input attributes should be eliminated from the sample.

Acknowledgement

The authors would like to thank OSeven Telematics, for providing all necessary data to accomplish this study.

References

Aarts, L., Van Schagen, I., 2006. Driving speed and the risk of road crashes: a review. Accid. Anal. Prev. 38 (2), 215–224. https://doi.org/10.1016/j.aap.2005.07.004.
Alper, D., Sinuany-Stern, Z., Shinar, D., 2015. Evaluating the efficiency of local municipalities in providing traffic safety using the Data Envelopment Analysis. Accid. Anal. Prev. 78, 39–50. https://doi.org/10.1016/j.aap.2015.02.014.
Andersson, Gunnar, Nilsson, Göran, 1997. Speed Management in Sweden: Speed, Speed Limits and Safety. Swedish National Road and Transport Research Institute.
Babaee, S., Bagherikahvarin, M., Sarrazin, R., Shen, Y., Hermans, E., 2015. Use of DEA and PROMETHEE II to assess the performance of older drivers. Transp. Res. Proc. 10, 798–808.
Babaee, S., Shen, Y., Hermans, E., Brijs, T., Wets, G., 2016. Measuring driver's relative performance over time using DEA and window analysis. In: Uncertainty Modelling in Knowledge Engineering and Decision Making: Proceedings of the 12th International FLINS Conference, pp. 885–890.
Babaee, S., Shen, Y., Hermans, E., Wets, G., Brijs, T., 2014a. Investigating individual driver performance: applying DEA on simulator data. In: Practical Applications of Intelligent Systems. Springer, Berlin, Heidelberg, pp. 623–635.
Babaee, S., Shen, Y., Hermans, E., Wets, G., Brijs, T., Ariën, C., 2014b. Combining driving performance information in an index score: a simulated curve-taking experiment. Transp. Res. Rec. 2434 (1), 44–51.
Backer-Grøndahl, Agathe, Sagberg, Fridulv, 2011. Driving and telephoning: relative accident risk when using hand-held and hands-free mobile phones. Saf. Sci. 49 (2), 324–330. https://doi.org/10.1016/j.ssci.2010.09.009.
Bejani, M.M., Ghatee, M., 2018. A context aware system for driving style evaluation by an ensemble learning on smartphone sensors data. Transp. Res. Part C: Emerg. Technol. 89, 303–320.
Birrell, Stewart A., Fowkes, Mark, Jennings, Paul A., 2014. Effect of using an in-vehicle smart driving aid on real-world driver performance. IEEE Trans. Intell. Transp. Syst. 15 (4), 1801–1810. https://doi.org/10.1109/TITS.2014.2328357.
Bonsall, Peter, Liu, Ronghui, Young, William, 2005. Modelling safety-related driving behaviour—impact of parameter values. Transp. Res. Part A: Policy Pract. 39 (5), 425–444. https://doi.org/10.1016/j.tra.2005.02.002.
Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. Eur. J. Oper. Res. 2 (6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8.
Cook, W.D., Seiford, L.M., 2009. Data envelopment analysis (DEA)–thirty years on. Eur. J. Oper. Res. 192 (1), 1–17. https://doi.org/10.1016/j.ejor.2008.01.032.
Cooper, W.W., Seiford, L.M., Tone, K., 2006. Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software. Springer Science & Business Media.
Desmond, Paula, Hancock, Peter, Monette, Janelle, 1998. Fatigue and automation-induced impairments in simulated driving performance. Transp. Res. Rec. J. Transp. Res. Board No. 1628 8–14. https://doi.org/10.3141/1628-02.
Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S., Shale, E.A., 2001. Pitfalls and protocols in DEA. Eur. J. Oper. Res. 132 (2), 245–259. https://doi.org/10.1016/S0377-2217(00)00149-1.
Eboli, L., Mazzulla, G., Pungillo, G., 2016. Combining speed and acceleration to define car users' safe or unsafe driving behaviour. Transp. Res. Part C: Emerg. Technol. 68, 113–125.
Eftekhari, H.R., Ghatee, M., 2016. An inference engine for smartphones to preprocess data and detect stationary and transportation modes. Transp. Res. Part C: Emerg. Technol. 69, 313–327.
Egilmez, G., McAvoy, D., 2013. Benchmarking road safety of US states: a DEA-based Malmquist productivity index approach. Accid. Anal. Prev. 53, 55–64. https://doi.org/10.1016/j.aap.2012.12.038.
Emrouznejad, A., Parker, B.R., Tavares, G., 2008. Evaluation of research in efficiency and productivity: a survey and analysis of the first 30 years of scholarly literature in DEA. Socio-Econ. Plan. Sci. 42 (3), 151–157. https://doi.org/10.1016/j.seps.2007.07.002.
Gadziński, J., 2018. Perspectives of the use of smartphones in travel behaviour studies: findings from a literature review and a pilot study. Transp. Res. Part C: Emerg. Technol. 88, 74–86.
Handel, Peter, Skog, Isaac, Wahlstrom, Johan, Bonawiede, Farid, Welch, Richard, Ohlsson, Jens, Ohlsson, Martin, 2014. Insurance telematics: Opportunities and challenges with the smartphone solution. IEEE Intell. Transp. Syst. Mag. 6 (4), 57–70. https://doi.org/10.1109/MITS.2014.2343262.
Haque, Md. Mazharul, Washington, Simon, 2015. The impact of mobile phone distraction on the braking behaviour of young drivers: a hazard-based duration model. Transp. Res. Part C: Emerg. Technol. 50, 13–27. https://linkinghub.elsevier.com/retrieve/pii/S0968090X14002137https://doi.org/10.1016/j.trc.2014.07.011.
Hollingsworth, B., Dawson, P.J., Maniadakis, N., 1999. Efficiency measurement of health care: a review of non-parametric methods and applications. Health Care Manage. Sci. 2 (3), 161–172. https://doi.org/10.1023/A:1019087828488.
Huang, H., Cheng, Y., Weibel, R., 2019. Transport mode detection based on mobile phone network data: a systematic review. Transport. Res. Part C: Emerg. Technol. 101, 297–312. https://doi.org/10.1016/j.trc.2019.02.008.
Jia, S., Hui, F., Li, S., Zhao, X., Khattak, A.J., 2019. Long short-term memory and convolutional neural network for abnormal driving behaviour recognition. IET Intell. Transp. Syst. https://doi.org/10.1049/iet-its.2019.0200. Print ISSN: 1751-956X and Online ISSN: 1751-9578.
Johnson, Derick A., Trivedi, Mohan M., 2011. Driving style recognition using a smartphone as a sensor platform. In: 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1609–1615. https://doi.org/10.1109/ITSC.2011.6083078.
Kanarachos, S., Christopoulos, S.R.G., Chroneos, A., 2018. Smartphones as an integrated platform for monitoring driver behaviour: the role of sensor fusion and connectivity. Transp. Res. Part C: Emerg. Technol. 95, 867–882.
Karlaftis, M.G., Gleason, J.M., Barnum, D.T. Bibliography of Urban Transit Data Envelopment Analysis (DEA) Publications. Available at SSRN: http://ssrn.com/abstract=1350583, https://doi.org/10.2139/ssrn.1350583, 2013.
Lenné, Michael G., Triggs, Thomas J., Redman, Jennifer R., 1997. Time of day variations in driving performance. Accid. Anal. Prev. 29 (4), 431–437. https://doi.org/10.1016/S0001-4575(97)00022-5.
Farrell, M.J., 1957. The measurement of productive efficiency. J. Roy. Stat. Soc. Ser. A 120 (3), 253–281. https://doi.org/10.2307/2343100.
Mantouka, E.G., Barmpounakis, E.N., Vlahogianni, E.I., 2018. Mobile Sensing and Machine Learning for Identifying Driving Safety Profiles (No. 18-01416).
Martić, Milan, Novaković, Marina, Baggia, Alenka, 2009. Data envelopment analysis-basic models and their utilization. Organizacija 42 (2), 37–43. https://doi.org/10.2478/v10051-009-0001-6.
Matthews, Gerald, Dorn, Lisa, Hoyes, Thomas W., Roy Davies, D., Ian Glendon, A., Taylor, Ray G., 1998. Driver stress and performance on a driving simulator. Hum. Fact. 40 (1), 136–149. https://doi.org/10.1518/001872098779480569.
Matthews, Gerald, Sparkes, Timothy J., Bygrave, Helen M., 1996. Attentional overload, stress, and simulate driving performance. Human Perform. 9 (1), 77–101. https://doi.org/10.1207/s15327043hup0901_5.
Mesken, Jolieke, Lajunen, Timo, Summala, Heikki, 2002. Interpersonal violations, speeding violations and their relation to accident involvement in Finland. Ergonomics 45 (7), 469–483. https://www.tandfonline.com/doi/full/10.1080/00140130210129682https://doi.org/10.1080/00140130210129682.
Nilsson, G., 1982. The Effects of Speed Limits on Traffic Accidents in Sweden. Swedish National Road and Transport Research Institute, Sartryck.
Nilsson, G., 2004. Traffic Safety Dimensions and the Power Model to Describe the Effect of Speed on Safety. Bulletin 221, Lund Institute of Technology, Lund. http://lup.lub.lu.se/record/21612.
Orfila, O., Saint Pierre, G., Messias, M., 2015. An android based ecodriving assistance system to improve safety and efficiency of internal combustion engine passenger

cars. Transp. Res. Part C: Emerg. Technol. 58, 772–782.
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. 119, 91–97. https://doi.org/10.1016/j.ssci.2019.05.059.
Patten, Christopher JD, Kircher, Albert, Östlund, Joakim, Nilsson, Lena, 2004. Using mobile telephones: cognitive workload and attention resource allocation. Accid. Anal. Prev. 36 (3), 341–350. https://doi.org/10.1016/S0001-4575(03)00014-9.
Shone, R., 1981. Applications in Intermediate Microeconomics. Martin Robertson, Oxford.
Ramanathan, Ramu, 2003. An Introduction to Data Envelopment Analysis: A Tool for Performance Measurement. Sage.
Saifuzzaman, Mohammad, Haque, Md Mazharul, Zheng, Zuduo, Washington, Simon, 2015. Impact of mobile phone use on car-following behaviour of young drivers. Accid. Anal. Prev. 82, 10–19. https://doi.org/10.1016/j.aap.2015.05.001.
Seiford, Lawrence M., 1997. A bibliography for data envelopment analysis (1978-1996). Ann. Operat. Res. 73, 393–438. https://doi.org/10.1023/A: 1018949800069.
Shichrur, Rachel, Sarid, Adi, Ratzon, Navah Z., 2014. Determining the sampling time frame for in-vehicle data recorder measurement in assessing drivers. Transp. Res. Part C: Emerg. Technol. 42, 99–106. https://doi.org/10.1016/j.trc.2014.02.017.
Speed management, 2006. OECD/ECMT Transport Research Centre (JTRC), Paris.
Sun, W., Zhang, X., Zhang, X., He, X., Zhang, G., 2019. Driving behaviour recognition based on orientation and position deviations. Int. J. Sens. Netw. 30 (3), 161–171.
Thanassoulis, Emmanuel, 2001. Introduction to the Theory and Application of Data Envelopment Analysis. Kluwer Academic Publishers, Dordrecht.
Toledo, Tomer, Musicant, Oren, Lotan, Tsippy, 2008. In-vehicle data recorders for monitoring and feedback on drivers' behavior. Transp. Res. Part C: Emerg. Technol. 16 (3), 320–331. https://doi.org/10.1016/j.trc.2008.01.001.
Törnros, Jan, Bolling, Anne, 2006. Mobile phone use–effects of conversation on mental workload and driving speed in rural and urban environments. Transp. Res. Part F: Traff. Psychol. Behav. 9 (4), 298–306. https://doi.org/10.1016/j.trf.2006.01.008.
Tselentis, I.D., Vlahogianni, I.E., George, Y., 2019. Hybrid data envelopment analysis for large-scale smartphone data modeling. World Conference on Transport Research.
Tselentis, Dimitrios I., Yannis, George, Vlahogianni, Eleni I., 2017. Innovative motor insurance schemes: a review of current practices and emerging challenges. Accid. Anal. Prev. 98, 139–148. https://doi.org/10.1016/j.aap.2016.10.006.
Vlahogianni, Eleni I., Barmpounakis, Emmanouil N., 2017. Driving analytics using smartphones: algorithms, comparisons and challenges. Transp. Res. Part C: Emerg. Technol. 79, 196–206. https://doi.org/10.1016/j.trc.2017.03.014.
Yannis, G., Papadimitriou, E., Papantoniou, P., 2014. Distracted driving and mobile phone use: overview of impacts and countermeasures. NTUA Road Safety Observatory. Proceedings of the Communication Technologies and Road Safety Conference, Abu Dhabi.
Young, K., Regan, M., 2007. Driver distraction: a review of the literature. In: Faulks, I.J., Regan, M., Stevenson, M., Brown, J., Porter, A., Irwin, J.D. (Eds.), Distracted Driving. Australasian College of Road Safety, Sydney, NSW, pp. 379–405.
Young, Mark S., Birrell, Stewart A., Stanton, Neville A., 2011. Safe driving in a green world: a review of driver performance benchmarks and technologies to support 'smart'driving. Appl. Ergon. 42 (4), 533–539. https://doi.org/10.1016/j.apergo.2010.08.012.
Zaldivar, Jorge, Calafate, Carlos T., Cano, Juan Carlos, Manzoni, Pietro, 2011. Providing accident detection in vehicular networks through OBD-II devices and Android-based smartphones. In: 2011 IEEE 36th Conference on Local Computer Networks (LCN). IEEE, pp. 813–819. https://doi.org/10.1109/LCN.2011.6115556.