# Star rating of driver's behavior as a tool to prevent risky behavior

Tešić, Milan; Miladić-Tešić, Suzana; Folla, Katerina; Yannis, George; Oviedo-Trespalacios, Oscar

**DOI**
10.1016/j.trf.2023.07.011

**Publication date**
2023

**Document Version**
Final published version

**Published in**
Transportation Research Part F: Traffic Psychology and Behaviour

**Citation (APA)**
Tešić, M., Miladić-Tešić, S., Folla, K., Yannis, G., & Oviedo-Trespalacios, O. (2023). Star rating of driver's behavior as a tool to prevent risky behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, *97*, 214-230. https://doi.org/10.1016/j.trf.2023.07.011

Transportation Research Part F: Psychology and Behaviour 97 (2023) 214–230

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

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

# Star rating of driver's behavior as a tool to prevent risky behavior

Milan Tešić [a,*], Suzana Miladić-Tešić [b], Katerina Folla [c], George Yannis [c], Oscar Oviedo-Trespalacios [d]

[a] The Road Traffic Safety Agency of the Republic of Serbia, Palace of Serbia, Blvd. Mihajlo Pupin 2, 11070 Belgrade, Serbia
[b] University of East Sarajevo, Faculty of Transport and Traffic Engineering, Vojvode Mišića 52, 74000 Doboj, Bosnia and Herzegovina
[c] Department of Transportation Planning and Engineering, National Technical University of Athens, Iroon Polytechniou 9, Zografou, 157 73 Athens, Greece
[d] Delft University of Technology, Faculty of Technology, Policy and Management, Section of Safety and Security Science, Jaffalaan 5, 2628 BX Delft, The Netherlands

ARTICLE INFO

*Keywords:*
Star rating
Distracted driving
Protective systems
Safety science
Safety management
Risk assessment

ABSTRACT

Driver behavior is a key determinant of road safety. Risky behaviors can be measured and quantified using traditional and modern methods, which enables the assessment of the driver's behavior. Following lessons from safety science, it is important to analyse and manage risks factors to optimize interventions in the transport context at the territory. The main objective of the paper is to explore the star rating of the driver's behavior to provide credible road safety monitoring and identification of the factors that most contribute to risky driving behaviors. The Data Envelopment Analysis method for selection of most significant factors is used for star rating of driver's behavior. The subject of the analysis encompasses 18 factors for 27 police administration units (PAUs) in the Republic of Serbia. The results are analysed regarding the following aspects: 1) star rating of driver's behavior and 2) defining the stages of the periodic factors monitoring for each PAU analysed. Using a mobile phone while driving, not using child restraints system, alcohol drunk driving during the night and seat belt use at rear seats of passenger vehicles were identified as the most significant factors associated with risky driving behaviors. Monitoring a broader set of factors helps to identify the strengths and weaknesses of a territory's road safety system. Star rating of driver's behavior as a tool intended for decision-makers ensures monitoring, management, the exchange of evidence-based and customized best practices and defining earlier goal-oriented actions to manage driver's behavior and to prevent risky driving.

## 1. Introduction

Globally, road traffic crashes cause nearly 1.3 million preventable deaths and an estimated 50 million injuries each year – making it the leading killer of children and young people worldwide (World Health Organization, 2022). Globally, road traffic crashes are the leading cause of death among children and young people aged 5–29 (World Health Organization, 2018). However, 20,678 people lost their lives on EU roads and around five more suffer serious injuries with life-changing consequences in 2022 (Carson et al., 2023). This is an unacceptable and unnecessary human and social price to pay for mobility. Moreover, progress in reducing road fatality rates EU-

* Corresponding author.
*E-mail addresses:* milan.te.sicm@gmail.com (M. Tešić), suzana.miladictesic@sf.ues.rs.ba (S. Miladić-Tešić), katfolla@central.ntua.gr (K. Folla), geyannis@central.ntua.gr (G. Yannis), o.oviedotrespalacios@tudelft.nl (O. Oviedo-Trespalacios).

https://doi.org/10.1016/j.trf.2023.07.011
Received 13 February 2023; Received in revised form 19 June 2023; Accepted 17 July 2023
Available online 26 July 2023
1369-8478/

wide has stagnated in recent years, excluding the year 2020, which was influenced by the COVID-19 pandemic. The development of road safety performance indicators has accelerated in the last decade. Efforts have been put in establishing similarities (and differences) among different countries worldwide. Today, modern methods of road safety comparisons encompass a multitude of factors (and consequently a multitude of indicators) while tending to reduce all those indicators to the same scale and allocate them as the most accurate weights possible to represent the specific features of the compared area. Depending on the purpose, the phase of selecting the representative road safety indicators on a territory should start from the analysis of all categories (levels) of indicators from the Koornstra et al. (2002) pyramid. The pyramid identifies four levels of (top-down) indicators, as follows: final outcomes (e.g., deaths per 100.000 inhabitants); intermediate outcomes (safety performance indicators); policy performance indicators (safety measures and programmes) and background performance indicators (structure and culture, i.e., road user's attitudes). Over the last couple of years, efforts have been made to identify correlations between certain "pyramid" levels, i.e., their influence on the final road safety assessment rate (Papadimitriou & Yannis, 2013). Given the variety of road safety indicators, it is necessary to identify the most important (significant) ones, providing a simple, yet realistic understanding of a road situation in a territory. Additionally, this can contribute to optimize the benefit of resources that are available for the implementation of road safety measures.

To better understand and identify road safety problems in its member-states, the European Commission (EC) continues with its efforts in improving road safety by defining the list of Key Performance Indicators (KPIs) within the *EU Road Safety Policy Framework 2021–20*30 (European Commission, 2019). The standardization of indicators from the point of view of defining and collecting them has been enabled in that way, helping decision-makers at all levels to monitor the progress towards achieving the „Vision Zero" by 2050. For all of these KPIs, the EC has defined a general methodological consideration applicable to all the indicators (see https://baseline.vias.be/). An overview of the availability gathered from 32 PIN countries (European Transport Safety Council, 2021) and *IRTAD Road Safety Annual Report 2020* (OECD/ITF, 2020), shows that there is still a way to go in terms of developing and collecting some of these KPIs. Most countries collect and analyze data related to safety belts (93.8 %), speed (87.5 %), alcohol (87.5 %), protective equipment (84.4 %), distraction (84.4 %) and vehicle safety (56.3 %). Only a few countries follow the KPIs related to infrastructure (34.4 %) and post-crash response (37.5 %). Half of these indicators are directly related to road user's behavior.

The core of the road safety problem lies in in-depth accident investigation studies, which provide a more complete picture of the real accident causes. The link between individual areas of the road safety system shows pivotal interaction between the vehicle, road and human behavior. "Human behavior factors" are the most accident contributing factors (93 %), while the second place is "road" (34 %) and "vehicle" (13 %), (PIARC, 2003). In accordance with these results, road user behavior is a key aspect of road safety and plays an important role in creating a clean, safe, and automated mobility ecosystem. The high-risk drivers and other road users are more likely to be involved in road crashes (Watson & Austin, 2021; Nguyen-Phuoc et al., 2020; Sheykhfard and Haghighi, 2020, 2019; Oviedo-Trespalacios & Scott-Parker, 2018; Scott-Parker & Oviedo-Trespalacios, 2017; etc.) and their behavior can be associated with a wide range of contributory factors, such as: alcohol-impaired driving, speeding (or driving above the speed limit), mobile phone use while driving, unfastened seat belt, etc.

Over the last years, digital technologies have been transforming the economy and society, affecting all sectors of activity, especially those of transport and mobility. Most of contributory factors can be measured and quantified using traditional (field or psychosocial research (Haghani et al., 2021, 2022)) and contemporary methods (unmanned aerial vehicle (Outay et al., 2020), data from automated vehicles and vehicle-to-everything communication-V2X (Khakzar et al., 2021; Ghosal & Conti, 2020; Kopelias et al., 2020; etc.), which enables the assessment of the driver's behavior and road safety performance level at the territory as well as to improve both road safety and efficiency by reducing the human driver variability in performance (Rahman and Abdel-Aty, 2021; Haque et al., 2021). Therefore, the automated process of KPIs data collection, accompanied by advanced smart solutions in urban areas, smart in-car solutions, etc. is expected in the near future.

Identifying road safety levels for cross-territory comparisons has attracted a lot of attention over recent years. The literature review has shown great efforts of the authors who tried to define the methodology for the assessment of the road safety level on a territory, including a wide range of indicators. The evolution of the idea of a road safety level assessment rate on territory has been transformed into two phases. Phase 1 encompasses the authors who made the calculations for the composite index based on indicators for only one layer (Jameel & Evdorides, 2021; Pires et al., 2020; Tešić et al., 2018; Bax et al., 2012; Shen et al., 2011; Hermans, 2009; Al-Haji, 2007 and etc.). Phase 2 gathers together the authors who calculated their composite index based on indicators of various layers (Babaee, 2022; Babaee et al., 2021; Chen et al., 2022, 2016; Hermans et al., 2010; Gitelman et al., 2010; Wegman and Oppe (2010); Wegman et al., 2008, 2005; etc.). The studies that have been conducted so far helped make the comparison of territories, define earlier goal-oriented road safety actions, and identify the best-in-class practices. The authors of these studies suggested creating a composite road safety index using the most appropriate indicators. This is not so simple because indicator-related data at international and national databases are not always available, and their definitions differ significantly. However, a composite road safety index obtained based on a broader comprehensive set of indicators provides a more accurate identification of good and poor road safety points for the territories concerned.

To assess the road safety level at the territory, the assessment, evaluation, and comparison of road safety factors related to vehicles and infrastructure have been performed using the Star Ratings Calculation Methodology. This methodology has been widely used nowadays in several transportation sectors, including 1) vehicle: New Car Assessment Programmes, i.e., GlobalNCAP, EuroNCAP, etc. and Green New Car Assessment Programmes; 2) road: Road Assessment Programme, i.e., iRAP, EuroRAP, etc. In order to improve road safety globally, star rating of the vehicle's safety features, or high-risk roads is carried out. In contrast, it was difficult to monitor and evaluate human behavior as the most important factor in road safety. Currently projects offering comparisons of territories using specific safety performance indicators, stand out the following: *Building the European Road Safety Observatory (SafetyNET project), Road Safety Data, Collection, Transfer and Analysis (DaCoTA project), ETSC Road Safety Performance Index (PIN program), Baseline project,* *Trendline project,* etc.

Applying the Star Ratings Methodology provides excellent promise for in-depth analysis and comprehension of human behavior with the goal of preventing risky driving. This is complicated because of the large number of risky behaviors that a driver can engage when driving and the difficulty studying these risky behaviors on the roads. However, a wide range of contributory factors can be combined in the star rating process. Given the variety of road safety factors, it is necessary to identify the most important (significant) ones, providing a simple, yet realistic understanding of the road situation in a territory. The star rating of driver's behavior enables identification of the territories where drivers behave less/more safely, and the credible cross-territory comparison according to driver's behavior. From the practical point of view, the star rating as a tool in the hands of decision-makers enables the exchange of best-performing practices between territories that have contributed to improving factors associated with risky driving behaviors and allocation of limited resources. This is of particular importance for the territories having limited resources for road safety investments, such as low and middle-income countries.

The main objective of the paper is to explore a star rating (SR) based on driver's behavior to provide credible road safety monitoring and identification of the factors that most contribute to road safety. The methodology will be used to compare and deeply analyzed driver's behavior in the 27 police administration units (PAUs) in the Republic of Serbia. The remaining part of this manuscript is structured as follows: description of the methodology including data collection and selection of factors, weighting and aggregation concepts and procedure for calculation of the star rating is given in Section 2. The results of the star rating, correlative analysis, and stages of the periodic monitoring of the factors in police administration units are discussed in Section 3. Section 4 is reserved for discussion of the most important results. The future perspectives of road safety performance indicators ecosystem and star rating of driver's behavior are presented in Section 5. The manuscript closes with the conclusions and topics for further research and limitations.

## 2. Methodology

The star rating of driver's behavior was developed using the upgraded methodology for developing a composite road safety performance index with the limited number of indicators for cross-territory (smaller territory) comparisons, developed by Tešić (2018). The upgrade of the methodology lies in its universality in the following three directions: 1) spatial: it is possible to include a larger number of countries, regions, etc. (territories) by adding appropriate data; 2) temporal: it is possible to span more years (time period), and 3) quantitative: it is possible to include a larger number of or more specific indicators. The upgraded methodology (star rating methodology) enables the star rating of driver's behavior at the smaller territory and identification of the factors that most contribute to risky driving behaviors. The following are descriptions of the three main components of this methodology: collection and selection of factors, weighting and data aggregation and procedure for calculation of the star rating of driver's behavior.

### 2.1. Collection and selection of factors

The subject of the analysis includes 18 safety performance indicators/factors for 27 PAUs, related to the driver's behavior that contributes to the occurrence and severe injuries of road accidents. The measurement of safety performance indicators in the Republic of Serbia is carried out in accordance with the methodological requirements defined in the EU Baseline project (see https://baseline.

**Table 1**
Spearman's correlation coefficient among factors associated with risky driving behaviors and output indicators.

| Code | Safety performance indicators-factors | PRWNC | TRWNC | PR Injured | TR Injured |
|---|---|---|---|---|---|
| | % of seat belt use by drivers of PV[a] in U[b] | −0.104 | −0.116 | −0.215 | −0.265 |
| | % of seat belt use by drivers of PV (U and R)[b] | −0.107 | −0.110 | −0.219 | −0.256 |
| | % of seat belt use at FRONT seats in PV (U) | −0.292* | −0.315* | −0.319* | −0.321** |
| | % of seat belt use at FRONT seats in PV (R) | −0.259 | −0.280* | −0.275* | −0.303* |
| **1** | **% of seat belt use at FRONT seats in PV (U and R)** | −0.288* | **−0.334*** | −0.304* | −0.311* |
| **2** | **% of seat belt use at REAR seats in PV (U and R)** | −0.378 | **−0.413*** | −0.347 | −0.391 |
| **3** | **% of drivers under the influence of alcohol-day (U and R)** | 0.273 | 0.143 | **0.486*** | 0.398* |
| **4** | **% of drivers under the influence of alcohol-night (U and R)** | 0.290 | 0.192 | **0.620**** | 0.524* |
| **5** | **% of drivers using mobile phone while driving (U and R)** | 0.224 | 0.123 | **0.426*** | 0.338* |
| | Average speed of PV in U | 0.342* | 0.278* | 0.254 | 0.186 |
| | 85th percentile of speed of PV in U | 0.291* | 0.232 | 0.265 | 0.208 |
| | Standard speed deviation of PV in R | −0.397** | −0.305* | −0.387** | −0.221 |
| **6** | **% of drivers exceeding the speed limit in U** | **0.363**** | 0.293* | 0.240 | 0.156 |
| | % of drivers exceeding the speed limit by 10 km/h in U | 0.320* | 0.275* | 0.303* | 0.195 |
| | % of use of daytime running lamps by drivers of PV (U and R) | −0.115 | −0.146 | −0.263* | −0.245 |
| **7** | **% of children using child restraints up to 3 years old (U and R)** | −0.325* | **−0.424*** | −0.374* | −0.281* |
| **8** | **% of children using child restraints from 4 to 12 years old (U and R)** | −0.248* | **−0.324*** | −0.314* | −0.301* |
| | % of children using child restraints up to 12 years old (U and R) | −0.146* | −0.227* | −0.212* | −0.288* |

*Correlation is significant at the level of 0.05 (1-tailed).
**Correlation is significant at the level of 0.01 (1-tailed).
[a] PV – passenger vehicle.
[b] Urban area (U), Rural area (R).

vias.be/en/about-the-project/) as well as national, science-based methodology for measuring and monitoring safety performance indicators, adopted in 2013, and revised in 2017, by the Road Traffic Safety Agency of the Republic of Serbia (RTSA). However, in order to analyze and star rating driver's behavior, available indicators that exclusively refer to driver behavior were used, while indicators related to vehicle safety, infrastructure safety and post-crash response were excluded from further analysis. Data have been taken from the open database of the RTSA (available at the link: https://www.abs.gov.rs/gis-baza) and data for each factor chosen to belong to the period [2016–2020], respectively, time-series data of five years. All data are taken from the afore-mentioned database as an *xlsx file. Data were analyzed and compared, and quality control was performed. The incomplete data were analyzed separately and supplemented from the data owner.

The following indicators have been used as output (final) indicators:1) Public risk based on the weighted number of casualties (fatal and non-fatal injuries) per 10,000 inhabitants (abbr. PRWNC); 2) Traffic risk based on the weighted number of casualties per 10,000 vehicles (abbr. TRWNC); 3) Public risk based on the number of the injured per 10,000 inhabitants (abbr. PR_Injured); 4) Traffic risk based on the number of the injured, per 10,000 vehicles (abbr. TR_Injured). The PRWNC and TRWNC are based on the type of consequences of road accidents weighted by corresponding coefficients, depending on the level of injuries. The weighting of output indicators is done according to Kukić et al., 2013. The following weights have been used in this analysis: 1) Fatality – weight 85; 2) Seriously injured – weight 10; 3) Slightly injured – weight 1. Factors used for the star rating of the driver's behavior have been chosen on the basis of the correlation coefficient value, i.e., the factor having the highest correlation coefficient with any of the 4 output (final) indicators has been chosen for further analysis (Table 1).

Analyzing time-series data of only five years, the results have shown that only eight factors (coded from 1 to 8) have a linear dependence on a minimum of one direct indicator whose correlation coefficient is bigger than 0.300 at $p = 0.05$, which are shown in Table 1, Column 1. The strength of the linear dependence varies among these eight factors and is ranging from $r = 0.324$ to $r = 0.620$, which is a characteristic of a "weak" and "mean" correlation strength (according to Pallant, 2011 classification). This correlation strength is expected because only time-series data of five years were analyzed. If the data set were to be extended for a longer period of time (i.e., ten years), then a higher degree of correlation strength is to be expected. One has to consider that the selection of factors for further analysis, within one risk domain, has been made by identifying the factor which is more comprehensive – wise (union of a few subsets-derived factors), and/or is in a significant correlation with minimum of one selected final outcome. For example, for the "speed" domain, there are three factors identified as having a significant correlation with a minimum of one selected final outcome. Their significance and influence on the occurrence of road crashes is stressed in the national methodology. It is important to note that the national methodology separately defines factors related to speeding and speeding up to 10 km/h, in urban and rural areas. Likewise, the methodology has foreseen monitoring of individual factors that concern the seat belt use by drivers, by passengers sitting in front seats, passengers sitting in rear seats, and the use of children seats. Furthermore, during the selection of factors, it was taken into account whether the factor was directly measured in the field. For example, the factor "% of drivers exceeding the speed limit in U" is more comprehensive because it was directly measured in the field and represents the union of two subsets such as: factor "% of drivers exceeding the speed limit by 10 km/h in U", and the derived factor "85th percentile of speed". Further, they are more precise and have a greater impact on the final outcomes (road accidents).

Measurements of the selected factors were performed once per year. Only factor "% of drivers using mobile phone while driving (U and R)" was not measured in 2018 in the Republic of Serbia. Using interpolation, the value for 2018 was obtained and it was used for further analysis. The re-scaling data for each of the eight factors for the 27 PAUs (27 PAUs × 5 time periods = 135) are described by means of the summary statistics (Table 2). Various normalization techniques exist (Nardo et al., 2005). Rescaling method, as a one of the various normalization techniques, has been used to normalize the factors data, based on the results of Hermans, 2009. The difference between each raw value and the factor minimum has been divided by the range of the factor. Consequently, the worst or lowest factor value is transformed into zero while the best or highest factor value obtains a rescaled score of one. All factors belong to the same scale and are reduced to the same basis. In other words, a factor's high (respectively low) value means more (or less) road casualties. Within that context, the direction of the factors F_3, F_4, F_5 and F_6 have been changed (Table 2, i.e., actual measured minimal value of the factor "% of drivers under the influence of alcohol-day (U and R)" on the field is 1.5 %, while after the change of direction, the factor value is 98.5 %; actual measured minimal value of the factor "% of drivers using mobile phone while driving (U and R)" is 7.22 %, while after the change of direction, the factor value is 92.78 %, etc.).

**Table 2**
Summary statistics on the re-scaling data.

| F | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| # of cases | | 135 | 135 | 135 | 135 | 135 | 135 | 135 | 135 |
| Mean | | 80.17 | 14.09 | 99.59 | 99.05 | 96.56 | 50.82 | 59.98 | 40.18 |
| Median | | 81.36 | 13.80 | 99.60 | 99.10 | 96.69 | 45.80 | 61.70 | 36.57 |
| Mode | | 77.80 | 10.60 | 100.00 | 100.00 | 97.47 | 46.50 | 74.68 | 42.83 |
| Std. Deviation | | 7.76 | 7.38 | 0.30 | 0.49 | 1.42 | 21.94 | 11.02 | 16.28 |
| Variance | | 60.25 | 54.46 | 0.09 | 0.24 | 2.01 | 481.43 | 121.42 | 264.92 |
| Skewness | | −0.40 | 1.16 | −0.98 | −0.34 | −0.22 | 0.59 | −0.84 | 0.33 |
| Minimum | | 59.54 | 1.54 | 98.50 | 97.69 | 92.78 | 8.30 | 25.43 | 5.15 |
| Maximum | | 97.90 | 48.71 | 100.00 | 100.00 | 99.33 | 93.40 | 80.95 | 75.61 |
| Percentiles | 25 | 74.00 | 8.91 | 99.40 | 98.78 | 95.58 | 36.00 | 53.89 | 30.07 |
| | 75 | 86.40 | 18.79 | 99.80 | 99.36 | 97.47 | 66.20 | 67.92 | 52.42 |

The table shows the values concerning the central tendency and distribution. As for the factor of the protective system, it can be seen that on average in the 27 PAUs, there is a seat belt wearing rate in the front seats of cars and vans of 80.17 %. Moreover, the most frequent value in the data set is 77.80 %, while the median equals 81.36 %. A variance of 60.25 or a standard deviation of 7.76 implies a considerable variability of protective systems data. Further, there is a significant change in the values of chosen factors among the PAUs. The values of factors related to the: protective system (F_1 and F_2); drivers exceeding the speed limit (F_6) and child restraints use (F_7 and F_8) fluctuate the most. The following factor whose value fluctuates significantly is the „percentage of drivers exceeding the speed limit in urban areas", (minimum: 8.30, maximum: 93.40). Factors relating to "alcohol" and "mobile phone" have the lowest variations in values.

### 2.2. Weighting and data aggregation

After the selection of the set of factors has been completed, making a selection of a weight allocation method seems to be the most important factor. According to Hermans et al., 2008, the results show that the Data Envelopment Analysis (DEA) method best approaches the ranking of final outcomes based on the number of road fatalities per million inhabitants. The Data Envelopment Analysis, originally developed by Charnes et al. (1978), is a non-parametric mathematical optimization technique used to assess the relative efficiency of a homogeneous set of decision-making units (DMUs), on the basis of multiple inputs and multiple outputs. The degree of other DMUs' inefficiency can be measured on the basis of their distance from the frontier. For each territory, there can be obtained a star rating score between zero and one, with higher values indicating better relative performance. For more information on this technique, we refer to Cooper et al. (2004) and Charnes et al. (1994).

The Ordered Weighted Averaging (OWA) operators are used as an expert method for data aggregation. This method is introduced by Yager (1988) and represents one of the most useful aggregation operators for the star rating because it enables road safety stakeholders to agree on the linguistic formulation to use this aggregation method. This provides a higher degree of acceptability of results obtained, which opens the door to the identification of the factors that most contribute to risky driving behaviors and the definition of the earlier goal-oriented road safety actions. For an OWA weighting vector, the degree of orness is defined as shown in Eq. (1) (Hermans, 2009). A common aggregation operator includes maximum, minimum and arithmetic mean. The weighting vector of these operators is given as: 1) max: $\vec{w} = (1,0,..,0)$, considering only the best performance; 2) min: $\vec{w} = (0,..,0,1)$, considering only the worst performance, and 3) arithmetic mean: $\vec{w} = (1/n,1/n..,1/n)$, considering each performance equally. In other words, $w_i$ reflects the importance of the $l^{th}$ indicator. One of the methods for obtaining relevant OWA weights is the "orness" concept. The degree of orness is an important numerical characteristic of averaging aggregation functions. The degree of orness corresponds to the degree of optimism of a decision maker, by defining the boundary of acceptable (safe) or unacceptable (unsafe) driver's behavior. More specially, the decision maker states the linguistic quantifier Q which offers a fuzzy description of the portion of criteria required to be met by a good solution.

$$orness\left(\overrightarrow{w}\right) = \frac{1}{l-1}\sum_{i=1}^{l}(l-i)\overrightarrow{w_i} = \frac{1}{l-1}\sum_{i=1}^{l-1}(l/i)^{\alpha} \tag{1}$$

$$\overrightarrow{w_i} = Q\left(\frac{i}{l}\right) - Q\left(\frac{i-1}{l}\right) \text{ for } i = 1, \cdots, l$$

In terms of road safety, α represents the degree to which the occurrence of road fatalities depends on the magnitude of the eight factors (listed in Table 1). The value of α that is larger (smaller) than one implies that the worst (best) performances affect the outcome indicators more and therefore low (high) factor values are emphasized. For the propose of linguistic formulation, the online panel discussion was organized, which included 17 road safety experts with more than 15 years of experience, from national universities and representatives from the RTSA. A national perspective on the contribution of each risk domain to road safety was obtained as the experts originated from different regions of the Republic of Serbia. However, if one wants to avoid compensation between good and bad scores, this method is the most useful aggregation operator for the star rating case because it enables the experts/decision makers/stakeholders at the national level to agree on the linguistic formulation for the purpose of this aggregation method. This also provides a higher degree of acceptability of the results obtained, which opens the door to the definition of the earlier goal-oriented actions. Linguistic formulations have been defined using the following principles: 1) In case a PAU scores badly on more than a few factors (minimum of three or 40 % of the total number of factors), its final star rating score should be small; and 2) In case a PAU scores badly on a few factors (a maximum of three or 40 % of the total number of factors), its final star rating score should be between small and average. The linguistic formulation of the threshold between safe and unsafe behavior represents at the same time a limitation of this analysis. This limitation is most pronounced in low- and medium-developed countries, because the threshold between safe and unsafe behavior depends on the level of dedication, knowledge, road safety culture of decision-makers, etc.

The first step in transforming the guidelines into restrictions for α is to give a specific meaning to the concepts 'badly' (with respect to factor) 'a few' (with respect to the number of factors), 'small' and 'average' (with respect to the star rating). As for the performance, it will be classified as 'good', 'average' or 'bad', depending on specific factors. Here, score 1 is assigned to good; score 0.5 to average and score 0 to bad performances. A 'small' index score is 0.25 at the most, an 'average' index score corresponds to 0.5 whereas a 'large' index score is at least 0.75. Therefore, based on (Eq. (1) and linguistic formulations, it was calculated that α should range in the interval [2.00; 2.950] in order to aggregate the eight factors in a way that is acceptable for the experts. The orness value in the interval [0.223; 0.313] is obtained by inserting the limit values of $\alpha$ in (Eq. (1). The result is the value of $\alpha$ equal to 2.0 and the OWA vector of (0.02;

0.05; 0.08; 0.11; 0.14; 0.17; 0.20; 0.23). The (Eq. (2) represents the algebraic model used to compute the star rating score for a *police administration unit j (j = 1…, n)*:

$$RSPI_j = \frac{\max}{\overline{w_{ij}}} \sum_{i=1}^{l} \overline{r_{ij} w_{ij}} \tag{2}$$

Subject to:

$$\sum_{i=1}^{l} \overline{w_{ij}} = 1$$

$$0.223 \leq \frac{1}{l-1} \sum_{i=1}^{l} (l-1)\overline{w_{ij}} \leq 0.313$$

$$L_m \leq \frac{r_{mj} w_{mj}}{\sum_{i=1}^{l} \overline{r_{ij} w_{ij}}} \leq U_m; \ \overline{w_{ij}} \geq 0$$

with l = number of factors; ¯ = ordered value; r = rescaled value; w = weight; m = {% of seat belt use at FRONT seats in PV (U and R); % seat belt use at REAR seats in PV (U and R); % of drivers under the influence of alcohol-day (U and R); % of drivers under the influence of alcohol-night (U and R); % of drivers using a mobile phone while driving; % of drivers exceeding the speed limit in U; % of children using child restraints up to 3 years old (U and R); % of children using child restraints from 4 to 12 years old (U and R)}; L = lower limit; U = upper limit.

The share of each of the eight factors in the overall star rating score is restricted by a lower and upper limit, for each factor. The limits have been defined using the budget allocation method, which included 17 road safety experts (Table 3). According to the DEA model presented, the average value of the two highest weights assigned by experts is used as the upper limit. The lower limit is obtained analogously.

The weights obtained by calculating the star rating score are used for the identification of the factors that most contribute to risky driving behaviors for this study since the programme could not find a feasible solution. The reason for that is the reduction in the space for searching an optimum solution based on three or four indicators. The additive aggregation method has been used in these cases, which helps obtain the value of the product of the allocated weight and the normalized value of indicator (defined by Eq. (2) without the software retrieval of the possible solution. The mentioned aggregation method has been used following the recommendations of Nardo et al., 2005 and those of Pešić, 2012, who made a test in which the linear aggregation method, based on pre-defined criteria, scored the best result. A system of "factor combinations" with three, four and five factors has been designed. The formula (Eq. (3) serves to determine the total number of factor combinations for the identification 3, 4, and 5 most significant factors associated with risky driving behaviors per PAUs:

$$C_k^n = \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k \cdot (k-1) \cdots 1}, \ n \geqslant k \geqslant 0, \ (n,k) \in N \tag{3}$$

Based on the previous equations: 1) the star rating with the limited number of factors ($SR_{ln}^3$) is calculated for 56 combinations; 2) the value of $SR_{ln}^4$ is calculated for 70 combinations, and 3) the value of $SR_{ln}^5$ is calculated for 56 combinations.

## 2.3. Procedures

The star rating of the driver's behavior methodology is calculated based on the following steps:

- **Step_1:** A previously made Excel sheet with the values of the eight factors is imported into the programme. The correlation analysis has been made possible using the IBM SPSS v.20 programme;
- **Step_2:** The overall star rating is calculated for all 27 PAUs. The system of combinations, optimization method and calculation of the star rating have been developed using the IBM CPLEX programme;
- **Step_3:** The system of "factor combinations" has been created accordingly, along with the calculation of values of $SR_{ln}^3$, $SR_{ln}^4$ and $SR_{ln}^5$;
- **Step_4:** A correlation analysis helped determine the correlation level between the SR and $SR_{ln}^n$;
- **Step_5:** Based on the SR correlation coefficients, the 3, 4 and 5 most significant factors associated with risky driving behaviors have been identified;
- **Step_6:** The values of $SR_{ln}^3$, $SR_{ln}^4$ and $SR_{ln}^5$ have been calculated for each PAU and

**Table 3**
Lower and upper limits for the share of each factor in the overall input.

| Code | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Upper limit | 0.1200 | 0.1125 | 0.0625 | 0.1175 | 0.0900 | 0.0700 | 0.1175 | 0.0250 |
| Lower limit | 0.2500 | 0.2150 | 0.1900 | 0.1900 | 0.1550 | 0.2000 | 0.2100 | 0.1550 |

- **Step_7:** Ranking of PAUs and defining the factors monitoring phases has been also made.

Following the previous steps, a list has been made containing the most significant factors associated with risky driving behaviors per PAUs that should be regularly monitored and combined into a high-quality star rating score. In order to better understand and interpret the results, the following definitions have been created for this study:

- PAU is an organizational unit of the Directorate of the Police of the Republic of Serbia outside the headquarters, which performs road safety tasks within its jurisdiction in a certain territory.
- Star rating with the limited number of factors (abbr. $SR^{n}_{ln}$) is a weighted combination of a limited number of factors,
- Star rating with the limited number (three/four/five) of factors (abbr. $SR^{3}_{ln}/SR^{4}_{ln}/SR^{5}_{ln}$) is a combination of a limited number of three/four/five factors and
- The most significant factors associated with risky driving behaviors are a set of indicators which, when calculating the star rating with the limited number of factors ($SR^{n}_{ln}$), results in the largest correlation with the overall star rating (SR) in only one police administration unit.

Finally, the methodology for star rating of driver's behavior is a tool, intended for researchers and decision makers, that enables the exchange of best-performing practices between territories that have contributed to improving factors associated with risky driving behaviors, allocation of limited resources intended for the improvement of road safety, define earlier goal-oriented road safety actions (preventive and enforcement activities aimed at drivers) for more monitoring and management of the driver's behavior, identify the

**Table 4**
Overall Star Rating results.

Overall Star Rating results

| 2016 | | 2017 | | 2018 | | 2019 | | 2020 | |
|---|---|---|---|---|---|---|---|---|---|
| **PAUs** | **SR** | **PAU** | **SR** | **PAU** | **SR** | **PAU** | **SR** | **PAU** | **SR** |
| Zaječar | 0.500 | Zaječar | 0.689 | Pančevo | 0.542 | Zaječar | 0.553 | Pirot | 0.639 |
| Prokuplje | 0.495 | Pirot | 0.590 | Belgrade | 0.526 | Sombor | 0.549 | Zaječar | 0.627 |
| Novi Sad | 0.443 | Čačak | 0.515 | Zaječar | 0.520 | Prokuplje | 0.513 | Pančevo | 0.567 |
| Belgrade | 0.417 | Subotica | 0.496 | Užice | 0.517 | Subotica | 0.504 | Sr. Mitrovica | 0.512 |
| Niš | 0.413 | Kikinda | 0.491 | Subotica | 0.513 | Pirot | 0.502 | Leskovac | 0.486 |
| Leskovac | 0.403 | Sombor | 0.471 | Prokuplje | 0.473 | Bor | 0.502 | Kruševac | 0.473 |
| Pančevo | 0.400 | Užice | 0.467 | Jagodina | 0.463 | Jagodina | 0.467 | Belgrade | 0.472 |
| Čačak | 0.399 | Pančevo | 0.443 | Zrenjanin | 0.458 | Kikinda | 0.462 | Prijepolje | 0.465 |
| Požarevac | 0.398 | Požarevac | 0.438 | Valjevo | 0.456 | Kruševac | 0.432 | Kragujevac | 0.462 |
| Jagodina | 0.394 | Novi Sad | 0.429 | Vranje | 0.451 | Belgrade | 0.428 | Kikinda | 0.458 |
| Sombor | 0.392 | Kruševac | 0.427 | Kikinda | 0.449 | Vranje | 0.428 | Prokuplje | 0.444 |
| Kikinda | 0.386 | Kragujevac | 0.416 | Kragujevac | 0.441 | Novi Sad | 0.421 | Smederevo | 0.442 |
| Prijepolje | 0.375 | Bor | 0.411 | Leskovac | 0.419 | Smederevo | 0.390 | Požarevac | 0.440 |
| Subotica | 0.365 | Leskovac | 0.410 | Smederevo | 0.414 | Požarevac | 0.378 | Užice | 0.437 |
| Kruševac | 0.358 | Belgrade | 0.404 | Novi Sad | 0.400 | Leskovac | 0.376 | Zrenjanin | 0.404 |
| Šabac | 0.355 | Šabac | 0.384 | Sr. Mitrovica | 0.394 | Pančevo | 0.374 | Šabac | 0.403 |
| Vranje | 0.350 | Valjevo | 0.382 | Kruševac | 0.363 | Čačak | 0.367 | Čačak | 0.379 |
| Pirot | 0.344 | Sr. Mitrovica | 0.357 | Požarevac | 0.330 | Sr. Mitrovica | 0.344 | Novi Sad | 0.372 |
| Zrenjanin | 0.331 | Kraljevo | 0.349 | Prijepolje | 0.301 | Kragujevac | 0.334 | Valjevo | 0.341 |
| Kraljevo | 0.329 | Smederevo | 0.305 | Kraljevo | 0.300 | Valjevo | 0.325 | Subotica | 0.331 |
| Sr. Mitrovica | 0.322 | Niš | 0.302 | Novi Pazar | 0.177 | Prijepolje | 0.295 | Vranje | 0.297 |
| Kragujevac | 0.316 | Zrenjanin | 0.267 | Niš | 0.000 | Užice | 0.292 | Sombor | 0.294 |
| Novi Pazar | 0.267 | Prijepolje | 0.220 | Pirot | 0.000 | Zrenjanin | 0.275 | Kraljevo | 0.254 |
| Bor | 0.000 | Novi Pazar | 0.077 | Bor | 0.000 | Kraljevo | 0.215 | Jagodina | 0.000 |
| Smederevo | 0.000 | Jagodina | 0.000 | Šabac | 0.000 | Niš | 0.000 | Niš | 0.000 |
| Valjevo | 0.000 | Prokuplje | 0.000 | Čačak | 0.000 | Šabac | 0.000 | Bor | 0.000 |
| Užice | 0.000 | Vranje | 0.000 | Sombor | 0.000 | Novi Pazar | 0.000 | Novi Pazar | 0.000 |

critical factors leading to road accidents, cross-territory comparison and strengthen the proactive road safety management at the territory (both at the international and national level). This is of particular importance for the territories having limited resources for road safety investments, such as low and middle-income countries.

## 3. Results

Along with the star rating of the driver's behavior, this paper identifies the most significant factors associated with risky driving behaviors for 27 police administration units in Serbia. The most significant factors are based on the correlation strength between the overall star rating scores and $SR^{n}_{ln}$, ensuring monitoring and exchange of evidence-based and customized best practice guidelines as well as defining earlier goal-oriented road safety actions to manage driver's behavior and to prevent risky driving.

### 3.1. Star rating of the driver's behavior

The star rating assessment of the driver's behavior is performed for each year and the whole examined period. Depending on the overall star rating scores (Table 4), the PAUs have been assigned stars using the following principle: 1) 5-star (green) PAUs whose value is above 0.750 and where territory drivers behave most safely while driving; 2) 4-star (yellow) PAUs whose value is between 0.500 and 0.749; 3) 3-star (orange) PAUs whose value is between 0.313 and 0.499; 4) 2-star (red) PAUs whose value is between 0.156 and 0.312 and 3) 1-star (black) PAUs whose value is below 0.155 and where territory drivers behave most unsafely while driving. Based on linguistic formulation, calculation in Eq. (2) as well as experts' assessment, the boundary between safe (orange, yellow and green) and unsafe (red and black) driver's behavior was obtained (0.313).

It could be noticed that there is no 5-star PAUs for the observed five time periods (Table 4). A 4-star PAUs can be perceived in 2016-Zaječar (0.500); in 2017-Zaječar (0.689), Pirot (0.590) and Čačak (0.515); in 2018-Pančevo (0.542), Belgrade (0.526), Zaječar (0.520), Užice (0.517) and Subotica (0.513); in 2019-Zaječar (0.553), Sombor (0.549), Prokuplje (0.513), Subotica (0.504), Pirot (0.502) and Bor (0.502); in 2020-Pirot (0.639), Zaječar (0.627), Pančevo (0.567) and Sremska Mitrovica (0.512). All other PAUs have a lower value of the overall star rating score when observing all time periods. As it can be seen, 2017 accounts for a successful year in terms of road safety as certain PAUs have achieved a high level of overall star rating score (0.689) of driver's behavior. The larger the number of factors used for the calculation of star rating, the higher the matching rate of these values is. There are PAUs among the surveyed ones whose rank is stable throughout all time periods, i.e., it does not change significantly. Some of them include the following ones: Zaječar, Kikinda, Niš, Kraljevo, etc., while some PAUs changed their position (rank) by one or more places (Pirot, Jagodina, Belgrade, Požarevac, Novi Sad, Čačak, Novi Pazar, etc.). The remaining PAUs have been found to have a significant rank deviation and consequently, their positions have changed (by five or more places), for example, Prokuplje, Bor, Valjevo, etc. The biggest differences in the ranking have been found at the PAU of Niš (between 5th (2016) and 21st (2017)) PAU of Prokuplje (between 2nd (2016) and 26th (2017) place), PAU of Bor (between 6th (2019) and 26th (2020) place), PAU of Valjevo (between 26th (2016), 17th (2017) and 9th (2018) place), etc. The reason for such fluctuations in the rank of the PAU of Niš in 2016 and 2017 are the oscillations in the value of individual factors. For example, the PAU of Niš accounted for the increase in indicator's value of "3. % of drivers under the influence of alcohol-day in urban and rural areas" by 0.62 % and indicator "4. % of drivers under the influence of alcohol-night in urban and rural areas" by 0.57 %, while the indicators: "5. % of drivers exceeding the speed limit in urban area" accounted for a drop in the value of indicator from 64.00 % to 48.10 %. Unlike the PAUs that are characterized by significant fluctuations, there have been identified PAUs having the minimum standard rank deviation, such as PAU of Novi Pazar, which "preserved" its last positions (from 21 to 27) throughout all the analyzed time periods.

As shown in Table 5, no PAUs have been awarded 5 stars, and only several of PAUs were awarded 4 stars for the road safety level of their driver's behavior during all time periods. The results show that 3-star PAUs are most represented, observing all analyzed time periods. The largest number of 3-star PAUs was in 2016 (77.78 %), 2017 (59.26 %), 2018 (48.15 %), 2019 (51.85 %) and 2020 (59.26 %). Further, a large number of PAUs were assigned one or two stars. For example, in 2020, 11.11 % of PAUs were assigned 2-stars while 14.81 % PAUs gained only 1-star, and these are the most unsafe PAUs where drivers behave most risky (unsafe). Generally, the overall star rating scores, as a consequence of the rank of a large number of PAUs, fluctuates during the analyzed time periods, which is shown in Fig. 1.

The application of Spearman's correlation analysis has contributed to investigate the dependence among the overall star rating scores and $SR^{n}_{ln}$ during the all-time periods for the entire territory of the Republic of Serbia (27 PAUs × 5 time periods = 135 item).

**Table 5**
Share of n-star PAUs.

| Star rating | 2016 | | 2017 | | 2018 | | 2019 | | 2020 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | No. of PAU | % | No. of PAU | % | No. of PAU | % | No. of PAU | % | No. of PAU | % |
| 5-star PAUs | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 |
| 4-star PAUs | 1 | 3.70 | 3 | 11.11 | 5 | 18.52 | 6 | 22.22 | 4 | 14.81 |
| 3-star PAUs | 21 | 77.78 | 16 | 59.26 | 13 | 48.15 | 14 | 51.85 | 16 | 59.26 |
| 2-star PAUs | 1 | 3.70 | 4 | 14.81 | 3 | 11.11 | 4 | 14.81 | 3 | 11.11 |
| 1-star PAUs | 4 | 14.81 | 4 | 14.81 | 6 | 22.22 | 3 | 11.11 | 4 | 14.81 |
| Total | **27** | **100.00** | **27** | **100.00** | **27** | **100.00** | **27** | **100.00** | **27** | **100.00** |

These results are shown in Table 6. *Column 1* contains combinations of factors marked by F codes, sorted by the value of the correlation coefficient in *Column 2*. Based on the guidelines for interpreting the correlations (Pallant, 2011), it is possible to conclude that, amongst the values of the star rating with the limited number of factors ($SR_{ln}^{3}$), all the combinations with three factors have strong correlations ($r \geq 0.564$, $r^2 \geq 0.318$, $p = 0.01$). The following are the most significant factors: "% of drivers using a mobile phone while driving in urban and rural areas"; "% of children using child restraints up to 3 years old in urban and rural areas" and "% of children using child restraints from 4 to 12 years old in urban and rural areas", (abbr.: 5_7_8). They offer the strongest linear dependence with the overall star rating score ($r = 0.907$, $r^2 = 0.823$, $p = 0.01$).

Looking at the results of Spearman's correlation analysis shown in Table 6, all values of the r correlation coefficient are greater than 0.564 (for $SR_{ln}^{3}$) or 0.691 (for $SR_{ln}^{4}$) or 0.757 (for $SR_{ln}^{5}$). The effect size for correlation is created by squaring the r values. According to the Cohen (1992, 1988) the effect size is small if the value of r varies around 0.1, medium if r varies around 0.3, and large if r varies more than 0.5. Based on that, all r values and therefore the values of $r^2$ (coefficient of determination) in Table 6 are greater than 0.5 (i. e., r value for SR_5 7 8 is 0.907 ($r^2 = 0.823$); SR_4 5 7 8, $r = 0.925$ ($r^2 = 0.856$) and SR_2 4 5 7 8, $r = 0.941$ ($r^2 = 0.886$)). The large effect size means that a study finding has practical significance. Hence, the Spearman rho showed a significant strong association between overall star rating scores, and the star rating obtained based on three, four or five of the most significant factors ($SR_{ln}^{n}$).

When observing *Column 3* and *4*, it can be noticed that all factor combinations offer a high correlation of the values of $SR_{ln}^{4}$ and the overall star rating ($r \geq 0.691$, $r^2 \geq 0.478$, $p = 0.01$ for all combinations). Compared with the previous three most significant factors, the list of the four most significant factors has been broadened with the factor „% of drivers under the influence of alcohol during the night (in urban and rural areas)", (abbr.: 4_5_7_8). The correlation coefficient of these factors is $r = 0.925$, $r^2 = 0.856$, $p = 001$. Finally,

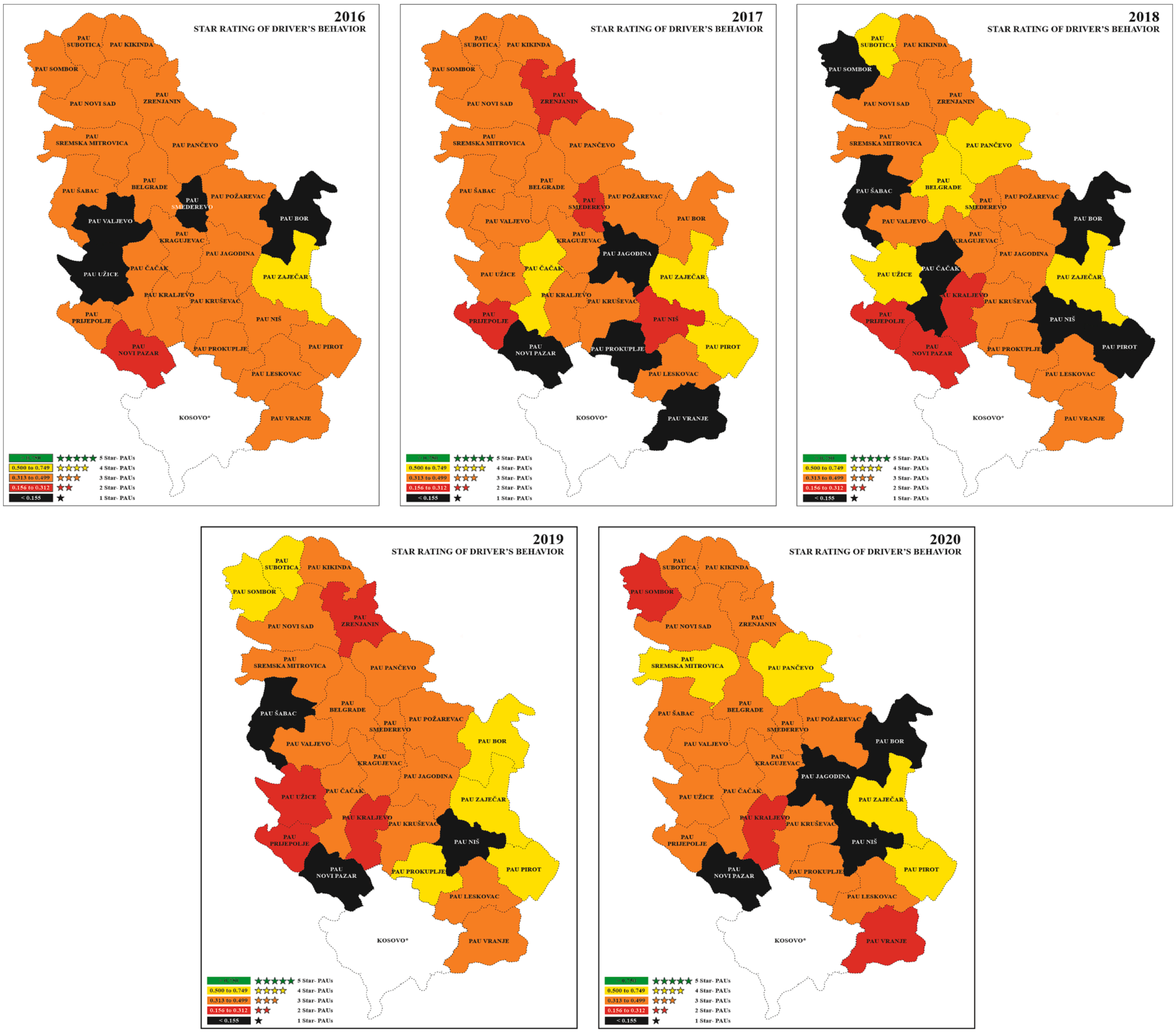


**Fig. 1.** Star rating of the driver's behavior, for 2016–2020, per year.

**Table 6**
The most significant factors based on Spearman's rho correlation coefficient with overall star rating scores, for the period 2016–2020.

| No of comb. | Column 1 | Column 2 | Column 3 | Column 4 | Column 5 | Column 6 |
|---|---|---|---|---|---|---|
| | $SR_{In}^{3}$ (F code) | Rank based on r values | $SR_{In}^{4}$ (F code) | Rank based on r values | $SR_{In}^{5}$ (F code) | Rank based on r values |
| 1. | SR_5 7 8 | 0.907* | SR_4 5 7 8 | 0.925* | SR_2 4 5 7 8 | 0.941* |
| 2. | SR_6 7 8 | 0.899* | SR_4 6 7 8 | 0.924* | SR_2 4 6 7 8 | 0.936* |
| 3. | SR_4 7 8 | 0.894* | SR_5 6 7 8 | 0.918* | SR_3 4 6 7 8 | 0.935* |
| 4. | SR_5 6 8 | 0.884* | SR_2 4 7 8 | 0.906* | SR_4 5 6 7 8 | 0.934* |
| 5. | SR_4 6 8 | 0.879* | SR_3 6 7 8 | 0.906* | SR_3 5 6 7 8 | 0.926* |
| . | . | . | . | . | . | . |
| 54. | SR_1 2 5 | 0.652* | SR_1 3 5 6 | 0.786* | SR_2 3 4 5 6 | 0.798* |
| 55. | SR_1 3 4 | 0.628* | SR_1 2 5 6 | 0.785* | SR_1 2 3 4 8 | 0.797* |
| 56. | SR_1 2 3 | 0.564* | SR_1 3 4 6 | 0.781* | SR_1 2 3 4 5 | 0.757* |
| . | | | . | . | | |
| 68. | | | SR_1 3 4 5 | 0.714* | | |
| 69. | | | SR_1 2 3 5 | 0.693* | | |
| 70. | | | SR_1 2 3 4 | 0.691* | | |

* Correlation is significant at the 0.01 level (1-tailed).

*Column 5* and *6* contain the rank of values of $SR_{In}^{5}$ with the overall star rating. The correlation coefficient of five factor combinations is very high ($r \geq 0.757$, $r^2 \geq 0.573$), which is somehow expected. As above, the list of the five most significant factors has been broadened with the factors "% seat belt use at REAR seats of passenger vehicles in urban and rural areas", (abbr.: 2_4_5_7_8), where $r = 0.941$, $r^2 = 0.886$, $p = 001$ (almost a complete matching of values of $SR_{In}^{5}$ with the overall star rating).

No matter which time period is observed, and which combination of factors is used for the star rating, the linear dependence strength of the values of $SR_{In}^{n}$ and the overall star rating scores belongs to the category of "strong positive relations". A high degree of matching of PAUs based on the values for the overall star rating and star rating with the three/four/five most significant factors is expressed through the values of the correlation coefficients in all time periods (Table 7). High correlation values range from $r = 0.713$, $r^2 = 0.508$, $p = 001$ in 2016 to $r = 0.972$, $r^2 = 0.945$, $p = 001$ in 2018.

The Wilcoxon Signed Ranks Test (Table 8) shows that there is no statistically significant change (the value of Asymp. Sig. (2-tailed) is higher than 0.05 in all cases) among the ranks of PAUs and the star rating with a limited number of factors have proved reliable for the national cross-territory comparisons, and for defining timely measures for the improvement of a road safety situation.

The results presented show that there is a strong linear dependence between the values of the overall star rating and $SR_{In}^{n}$, and that the applied methodology for star rating of driver's behavior provides a reliable and credible national cross-territory comparison. Also, a strong, positive linear dependence of the ranks of PAUs, ranked according to various factors combinations, confirms the reliability of the methodology applied.

### 3.2. Identification of the most significant factors associated with risky driving behaviors per PAUs

Large correlations identified between the calculated overall star rating and star rating with a limited number of factors associated with risky driving behaviors show that it is possible to describe the road safety level of driver's behavior and compare the PAUs, reliably. The matching of PAU' rankings in the function of a change of factors included in the star rating is considerable. Apart from factor selection contributing most to the reliability, comparability, credibility, and sensitivity of a star rating, change in factors' values over time (through several time periods) influences significantly the sensitivity of the star rating. One of the essential problems that can lead to changes in star rating' scores within a specific time period, i.e., over several time periods, is the lack of a periodical, clearly defined and sustainable national measuring system. Therefore, decision-makers have no possibility to precisely define their priorities. As a consequence, the measures undertaken may not be sufficiently effective.

**Table 7**
Correlation analysis between the overall star rating and star rating with the most significant factors in the time period.

| | | | | | | |
|---|---|---|---|---|---|---|
| **2016** | | | | **$SR_{In}^{3}$: 2 6 8** | **$SR_{In}^{4}$: 2 6 7 8** | **$SR_{In}^{5}$: 2 4 6 7 8** |
| | Spearman's rho | **SR** | Correlation Coefficient | 0.713** | 0.718** | 0.756** |
| **2017** | | | | **$SR_{In}^{3}$: 3 6 8** | **$SR_{In}^{4}$: 2 3 6 8** | **$SR_{In}^{5}$: 2 3 4 6 8** |
| | Spearman's rho | **SR** | Correlation Coefficient | 0.961** | 0.964** | 0.965** |
| **2018** | | | | **$SR_{In}^{3}$: 6 7 8** | **$SR_{In}^{4}$: 4 6 7 8** | **$SR_{In}^{5}$: 1 4 6 7 8** |
| | Spearman's rho | **SR** | Correlation Coefficient | 0.969** | 0.970** | 0.972** |
| **2019** | | | | **$SR_{In}^{3}$: 5 6 8** | **$SR_{In}^{4}$: 5 6 7 8** | **$SR_{In}^{5}$: 2 3 6 7 8** |
| | Spearman's rho | **SR** | Correlation Coefficient | 0.933** | 0.954** | 0.966** |
| **2020** | | | | **$SR_{In}^{3}$: 6 7 8** | **$SR_{In}^{4}$: 3 6 7 8** | **$SR_{In}^{5}$: 3 4 6 7 8** |
| | Spearman's rho | **SR** | Correlation Coefficient | 0.942** | 0.949** | 0.970** |
| | | | Sig. (1-tailed) | 0 | 0 | 0 |
| | | | N | 27 | 27 | 27 |

** Correlation is significant at the level of 0.01 (1-tailed).

**Table 8**
Wilcoxon Signed Ranks Test of the rankings in the period 2016–2020 (in total, five time periods).

| Wilcoxon Signed Ranks Test Statistics | | The most significant factors | | |
|---|---|---|---|---|
| | | $SR^3_{ln}$: 5 7 8 - SR | $SR^4_{ln}$: 4 5 7 8 - SR | $SR^5_{ln}$: 2 4 5 7 8 - SR |
| 2016–2020 | Z | −.433[a] | −.385[b] | −.056[b] |
| | Asymp. Sig. (2-tailed) | 0.665 | 0.700 | 0.955 |

[a] Based on positive ranks.
[b] Based on negative ranks.

**Table 9**
The most significant factors per PAUs.

| | $SR_{ln}^3$ | | | | $SR_{ln}^4$ | | $SR_{ln}^5$ | |
|---|---|---|---|---|---|---|---|---|
| Police administration unit | Three the most significant factors (% of SR score) | | | R value | 4th factor (% of SR score) | R value | 5th factor (% of SR score) | R value |
| Belgrade | 2 | 3 | 8 (31.61) | 0.999 | 1 (46.81) | 0.995 | 4 (63.73) | 0.996 |
| Kragujevac | 1 | 6 | 7 (39.73) | 0.980 | 8 (47.37) | 0.989 | 3 (59.46) | 0.992 |
| Jagodina | 1 | 2 | 3 (27.75) | 0.980 | 4 (38.58) | 0.984 | 5 (53.66) | 0.981 |
| Niš | 1 | 2 | 4 (40.32) | 0.984 | 3 (50.59) | 0.990 | 6 (57.52) | 0.996 |
| Pirot | 1 | 2 | 8 (31.96) | 1.000 | 3 (42.28) | 1.000 | 5 (56.21) | 1.000 |
| Prokuplje | 1 | 2 | 3 (29.75) | 0.973 | 7 (45.06) | 0.978 | 5 (57.31) | 0.967 |
| Leskovac | 1 | 2 | 4 (30.63) | 0.996 | 8 (40.96) | 0.997 | 7 (51.60) | 0.998 |
| Vranje | 1 | 3 | 5 (39.02) | 0.996 | 2(50.75) | 0.997 | 4 (61.68) | 1.000 |
| Zaječar | 3 | 5 | 7 (41.24) | 0.937 | 1 (49.58) | 0.951 | 2 (59.50) | 0.983 |
| Bor | 1 | 2 | 3 (32.99) | 0.995 | 4 (48.73) | 0.995 | 5 (60.60) | 0.997 |
| Smederevo | 3 | 6 | 7 (39.52) | 0.938 | 5 (53.51) | 0.960 | 1 (65.51) | 0.970 |
| Požarevac | 2 | 5 | 6 (37.10) | 0.988 | 1 (50.94) | 0.993 | 7 (65.83) | 0.998 |
| Valjevo | 1 | 4 | 7 (41.17) | 0.990 | 5 (50.93) | 0.997 | 6 (64.11) | 0.989 |
| Šabac | 5 | 6 | 8 (33.61) | 0.997 | 4 (47.67) | 0.995 | 3 (57.60) | 0.997 |
| Kraljevo | 5 | 6 | 8 (35.06) | 0.995 | 1 (47.57) | 0.996 | 3 (62.41) | 0.998 |
| Kruševac | 2 | 6 | 8 (45.75) | 0.979 | 4 (50.38) | 0.979 | 3 (62.99) | 0.991 |
| Čačak | 3 | 4 | 7 (40.68) | 0.970 | 2 (52.93) | 0.994 | 5 (63.77) | 0.990 |
| Novi Pazar | 1 | 2 | 3 (51.48) | 0.993 | 4 (68.51) | 0.998 | 5 (87.39) | 0.997 |
| Užice | 3 | 6 | 8 (35.64) | 1.000 | 7 (51.60) | 0.999 | 4 (62.55) | 1.000 |
| Prijepolje | 2 | 3 | 6 (40.56) | 0.995 | 8 (45.33) | 0.997 | 4 (58.22) | 0.997 |
| Novi Sad | 4 | 5 | 6 (42.40) | 0.973 | 8 (57.91) | 0.995 | 1 (67.30) | 0.989 |
| Zrenjanin | 2 | 6 | 7 (40.00) | 0.995 | 5 (49.85) | 0.999 | 8 (58.50) | 0.998 |
| Kikinda | 2 | 4 | 6 (34.72) | 0.996 | 7 (48.61) | 0.998 | 8 (62.97) | 1.000 |
| Pančevo | 2 | 5 | 8 (38.75) | 0.980 | 1 (51.70) | 0.989 | 7 (68.26) | 0.994 |
| Sombor | 1 | 2 | 4 (35.53) | 0.808 | 5 (42.30) | 0.864 | 8 (57.66) | 0.957 |
| Subotica | 2 | 4 | 7 (39.48) | 0.990 | 5 (51.92) | 0.999 | 8 (63.03) | 0.995 |
| Sremska Mitrovica | 4 | 5 | 7 (46.92) | 0.991 | 3 (63.27) | 0.991 | 6 (80.02) | 0.998 |
| 1 | *% of seat belt use at FRONT seats in PV (U and R)* | | | | | | | |
| 2 | *% seat belt use at REAR seats in PV (U and R)* | | | | | | | |
| 3 | *% of drivers under the influence of alcohol- day (U and R)* | | | | | | | |
| 4 | *% of drivers under the influence of alcohol- night (U and R)* | | | | | | | |
| 5 | *% of drivers using mobile phone while driving (U and R)* | | | | | | | |
| 6 | *% of drivers exceeding the speed limit in U* | | | | | | | |
| 7 | *% of children using child restraints up to 3 years old (U and R)* | | | | | | | |
| 8 | *% of children using child restraints from 4 to 12 years old (U and R)* | | | | | | | |

The influence of these factors differs from one territory to another. This means that the lists of the most significant factors in each PAUs differ significantly, and so does the share of various most contributing combinations of factors in the overall star rating score of each PAU. Based on the results of Spearman's correlation analysis, the most significant factors were identified. Table 9 gives an overview of the most significant factors and shows a visual display of a representation of specific factors among the most significant combination of factors per PAUs. For example, the most significant factors associated with risky driving behaviors in the PAU of Belgrade are: % of seat belt use at REAR seats in passenger vehicles (in urban and rural areas); % of drivers under the influence of alcohol during the day (in urban and rural areas) and % of children using child restraints from 4 to 12 years old (in urban and rural areas). They give the $SR^3_{In}$ the value of 2_3_8 = 0.142. This value represents 31.61 % of the value of the overall star rating obtained based on all eight factors for all time periods. This value is obtained as the mean value of the star rating with a limited number of factors for all time periods. With the increase in the number of factors used for the calculation, the share of the $SR^n_{In}$ is increasing in the overall star rating score. The first next factor yielding the highest value of the star rating score (PAU of Belgrade), as part of a combination, is the „% of seat belt use at FRONT seats in passenger vehicles (in urban and rural areas)" ($SR^4_{In}$: 1_2_3_8 = 0.210, while the share in the value of the overall star rating amounts to 46.81 %). Finally, the factor „% of drivers under the influence of alcohol the night (in urban and rural areas)" makes a combination of factors that gives the closest value of the $SR^5_{In}$: 1_2_3_4_8 = 0.287 (representing 63.73 % of the overall star rating score for the PAU of Belgrade). Similar analysis can be made for all analyzed police administration units, with the exception of those police administration units having the overall star rating score of 0 (these PAU performed worst in the optimization test).

The results show that there is a strong linear dependence between the overall star rating scores and values of $SR^n_{In}$, and that the applied methodology for star rating driver's behavior provides a reliable and credible national cross-territory comparison. When observing Column "r value", it can be noticed that all factor combinations offer a high correlation of the values of $SR^n_{In}$ and the overall star rating per PAU ($r \geq 0.808$, $r^2 \geq 0.653$, $p = 0.01$ for all combinations of three factors; $r \geq 0.951$, $r^2 \geq 0.904$, $p = 0.01$ for all combinations of four factors; and $r \geq 0.957$, $r^2 \geq 0.912$, $p = 0.01$ for all combinations of five factors). Furthermore, the results have shown that it is possible to identify the factors closely associated with risky driving behaviors for each police administration unit and define the stages of the periodical monitoring of the factors. The stages are defined based on the most significant factors. In the first stage, PAU should monitor the first three most significant factor (i.e., Belgrade should monitor factors 2, 3 and 8), then in the second stage, PAU should monitor next, fourth factors (i.e., Belgrade should monitor factor 1), and in the third stage, PAU should monitor fifth factor (i.e., Belgrade should monitor factor 4). Observing the results for all PAUs (Table 9), can be concluded that the most frequent factors in the first stage are "% of seat belt use at FRONT seats in PV (U and R)" which was identified in 11 PAUs, then "% seat belt use at REAR seats in PV (U and R)", and "% of children using child restraints up to 3 years old (U and R)", which was identified in 8 PAUs. In the second stage, factors that appear in the five PAUs are "% of drivers under the influence of alcohol-night (U and R)" and "% of drivers using mobile phone while driving (U and R)". Finally, factor "% of drivers using mobile phone while driving (U and R)" needs to be monitored in the third stage even at 6 PAU. In addition, one of the factors contributing to the variety of the most significant combinations of factors are the fluctuations in indicator values, both of time periods and police administration units.

## 4. Discussion

### 4.1. Star rating of driver's behavior at the national level

The star rating at the national level has been made using factors that are linearly dependent on a minimum of one output indicator. The results have shown that only eight factors have a linear dependence on a minimum of one output indicator. The results shows that the PAUs differ significantly from the point of view of driver's behavior (Fig. 1). The Spearman's correlation analysis of the results obtained by calculating the star rating for various combinations of three, four and five factors and the overall star rating has helped identify the most significant factors associated with risky driving behaviors on the Republic of Serbia. As given in Table 6 and Table 7 the correlation analysis has shown that all the combinations have a high correlation ($r \geq .564, r^2 \geq .318, p = .01$) with the overall star rating. Therefore, there is a high level of matching of ranks of PAUs regardless of which combination is taken for the calculation of the overall star rating. The greatest matching is with the values of $SR^5_{In}$ and the smallest one with the $SR^3_{In}$, as expected since the value of the overall star rating is more precise when the calculation comprises a larger number of factors. Table 8 shows that there is no statistically significant change among the ranks of PAUs. Based on the above-mentioned, the star rating of driver's behavior is a reliable and robust tool for defining road safety (i.e., preventive and enforcement) actions to improve the driver's behavior. Hence, the application of the methodology (approach) and obtained results are curial for in-depth understanding driver's behavior in low- and middle-income countries, such as Republic of Serbia and equally important, national cross-territory comparisons. The most significant factors associated with risky driving behaviors that have been identified in that way ensure the optimum choice of factors. The following are the most significant factors during the all-time periods for the entire territory of the Republic of Serbia:

- Three most significant factors: "% of drivers using a mobile phone while driving in urban and rural areas"; "% of children using child restraints up to 3 years old in urban and rural areas" and "% of children using child restraints from 4 to 12 years old", (abbr.: 5_7_8), ($r = .907, r^2 \geq .823, p = 0.01$);
- The fourth, most significant factor is „% of drivers under the influence of alcohol during the night (in urban and rural areas)", (abbr.: 4_5_7_8), ($r = .925, r^2 \geq .856, p = 0.01$); and
- The fifth most significant factor is "% seat belt use at REAR seats of passenger vehicles in urban and rural areas", (abbr.: 2_4_5_7_8), ($r = .941, r^2 \geq .886, p = 0.01$).

The obtained results reflect the road safety state in the Republic of Serbia. In the period from 2016 to 2020, the percentage of child road deaths who died as passengers among the total number of child road deaths in the Republic of Serbia ranged from 41.7% to 70 %, respectively, and the percentage of injured children as passengers ranged from 27 % up to 32.1 %. Furthermore, the share of drivers killed who were under the influence of alcohol in the total number of dead drivers in road accidents ranged from 11.7 % to 17.7 %. At the safety performance indicators level in the Republic of Serbia, the percentage of children using child restraints up to 3 years old ranged from 39.1 % to 69.2 %, while the percentage of children using child restraints from 4 to 12 years old ranged from 39.1 % to 69.2 %, which is very low compared to developed countries. Further, the value of the factor related to the seat belt use at rear seats in the passenger vehicle is significantly low and ranged from 10.1 % to 20.6 %, while the factor related to alcohol-impaired driving ranged from 0.80 % to 1.06 % and mobile phone use while driving ranged from 3.5 % to 5.6 %.These road safety data show that the star rating gives precise results that enable decision makers to act reliably and invest in improving road safety in territory under their jurisdiction.

From the practical point of view, the star rating of driver's behavior as a tool in the hands of the decision-maker enables them to guide preventive and enforcement activities according to the identified most significant factors. Based on the obtained results, road safety stakeholders in the Republic of Serbia should direct their activities towards the above-mentioned factors. It is necessary to carry out preventive activities (various campaigns on social networks, education, peer education among young people, etc.) which indicate the risk of using a mobile phone while driving, the risk of driving under the influence of alcohol, and the importance of seat belts as a passive element of road safety. The results, on the other hand, make it possible to choose a PAU quickly and precisely, directing actions in accordance with the road type (urban or rural roads or motorways) and the time of day (day or night), in which the Ministry of Interior should immediately step up its enforcement measures. Especially, if the efficiency rate of preventive and enforcement measures defined in the Safety Cube project (see https://www.roadsafety-dss.eu/#/) and Daniels et al., 2019 are considered. Namely, the results of the project show that preventive activities/campaigns related to the seat belt use have an efficiency rate ("cost/benefit" rate) of 42.2 (Aigner-Breuss, 2017), enforcement of seatbelt wearing for light-vehicle occupants have an efficiency rate of 17.7 (Daniels et al., 2019), use of children's car seats has an efficiency rate of 4.6 (Kaiser, 2017). If such road safety activities are carried out continuously, it is possible to reduce the number of deaths by up from 8 % to 50 %, or seriously injured people by up from 8 to 45 %. By conducting road safety activities in this way, it is possible to reduce the influence of "human behavior factors" on the occurrence of road accidents, which leads to the final reduction of road deaths and injuries.

At the national level, the star rating results show that no PAUs have been awarded 5 stars. Additionally, 3-star PAUs were the most common. Further, a large number of PAUs are assigned one or two stars, which is extremely concerning as it suggests a large prevalence of risky behavior. The fact that the overall star rating scores for numerous PAUs changed throughout the course of the observation period is an interesting finding. This variation is the result of a lack of systematic road safety management at the national level, meaning management that is not based on science, data, and clearly defined strategic targets and objectives. Indeed, road safety initiatives conducted during the study period did not represent the current state of road safety, highlighting the need for a better national road safety strategy and action plan.

### *4.2. Identification of the most significant factors associated with risky driving behaviors per PAUs*

The analysis of the comprehensive statistics of factors included in the star rating (Table 2) has shown that there is a significant standard deviation of factor values in the observed time periods. Apart from the standard deviation, the value of variance and the skewness has shown that there is a significant change in the value of selected factors among the observed PAUs. The statistics have also shown that PAUs in the Republic of Serbia differ significantly from the point of risky driving behavior, i.e., "alcohol-impaired driving", "speeding", "mobile phone" and "protective systems-seat belts and child restraints". This suggests that all the PAUs have not equally targeted risky driving behavior.

The existing variety among the PAUs provides guidance on the selection of the most significant factors that should be targeted on each territory, considering their specific features. This is particularly important as human, material and financial resources are limited in the territories and there is a need to optimize the number of factors. This will contribute to a more efficient road safety management of the PAUs, allowing road safety stakeholders to engage in preventive road safety actions.

The stages of periodic monitoring to improve road safety at the PAU are defined based on identified the most significant factors. The analysis of the results helped to observe the following PAUs, where one more factor is added in each next stage of factors monitoring. The concept of monitoring and managing a sufficient number of the most significant factors and the addition of the next most significant factor will help develop the periodical and sustainable system for managing factors associated with risky driving behaviors across PAUs. Also, this concept encourages as many territories as possible to develop the periodical and sustainable system for measuring and managing factors, as the methodology proposed provides the best efficient ratio of invested resources and obtained results, when it comes to their managing.

Based on the previously obtained results, it is possible to identify the most significant behavioral factors associated with the safety of drivers. The systematic monitoring of both the evolution of these behavioral factors and the road safety outcomes can contribute to define goal-oriented road safety actions and opt for the most appropriate activities in order to prevent risky driving in each PAU. Specifically, the results presented herein have shown that it is possible to identify a group of factors representing in a credible way the road safety level of driver's behavior in PAUs and provide their comparison with high precision. Therefore, the $SR_{In}^{n}$ is reliable and robust enough to enable timely defining of preventive measures for road safety system improvement and defining the factors monitoring phases to prevent risky driving for each PAU analyzed. Also, it enables the comparison of the largest possible number of territories (both at the international and national level) which contributes indirectly to setting up a system of factor monitoring on these territories. Recording a broader comprehensive set of factors associated with risky driving behaviors helps identify the strengths and weaknesses of a territory's road safety system.

Analysis of the level of road safety culture, i.e., attitudes and behavior of road users is becoming more and more relevant today (Meesmann et al., 2022; Babaee et al., 2021; etc.) The innovativeness of this approach is reflected in the in-depth assessment of behavior (road safety culture) of a certain, specific category of road users (in this case, drivers) in a territory. Compared to the previous analysis in Tešić et al., 2018, which analyzed 21 European countries and 6 safety performance indicators related to road, road users and post-crash response, star rating methodology included 18 specific indicators related to driver's behavior only, measured at smaller organizational units (PAUs) within one country. Finally, these results confirm the universality of the original methodology developed by Tešić, 2018, in the three directions: spatial, temporal, and quantitative.

## 5. Future perspectives of road safety performance indicators

The star rating of safety performance indicators (in addition to monitoring road user's attitudes) is one of the modern approaches having the highest potential for road safety improvement in the territory. One of the preconditions for the application of this modern approach is the standardization of safety performance indicators from the point of view of defining and collecting. A significant step towards the standardization of indicators at the EU level was made by the EC by defining the list of Key Performance Indicators (KPIs) within *EU Road Safety Policy Framework 2021–2030*. The next, very significant, precondition that contributes to the creation of a sustainable road safety performance indicators data ecosystem (RSPI data ecosystem) is the *Baseline* project (see https://baseline.vias.be/). Through this project, the European Commission has defined a general methodological consideration applicable to all KPIs.

Road safety decision-making requires open data and standardisation. One of the biggest initiatives to promote the Open Science in transport research is the project BE OPEN, which is funded by the EC within the H2020 project (see https://beopen-project.eu/the-project). Within this project, Yannis et al., 2020 are emphasized the importance of collecting road safety performance indicators as a part of the platform for global road safety data analysis. Following the development trends of open data platforms, the proposed star rating methodology can be used as an *open tool for road safety performance indicators monitoring and management* making an open RSPI data ecosystem. The star rating methodology as an open tool implies direct involvement of the EC Directorate-General for Mobility and Transport, as a focal point for KPIs management, at the EU level. This concept provides a strong collaboration with national road safety stakeholders which is responsible for KPIs monitoring at the national level, monitoring road safety progress, defining earlier goal-oriented road safety actions for road safety improvement, identifying the critical factors leading to road accidents, and strengthening proactive road safety management. From the practical point of view, the following future challenges need to be overcome:

- All Member States need to define a comprehensive methodology for collecting and monitoring KPIs at the national level, which is completely in line with the EC minimum methodological requirements;
- The national methodology should define a leading road safety stakeholder for collecting KPIs (most often it is a ministry of transport or leading road safety agency);
- The government/leading road safety stakeholder should establish a sustainable funding source for periodic, long-term monitoring of safety performance indicators;
- The leading road safety stakeholder should establish the mechanisms for reporting to the parliament, citizens, etc.

The limitations can occur in the semantic layer when it comes to differences in the definitions of certain indicators i.e. (different legal requirements concerning helmet use by cyclists, applicable legal provisions relating to the maximum permitted blood alcohol content, the definition of road safety performance indicators related to the vehicle and infrastructure). Also, the numerous limitations that can hinder the reuse of RSPIs data are listed in Böhm et al. (2018), among which the following ones stand out: data storage, fragmentation of data ownership, a lack of interoperability between datasets and platforms, etc.

However, automated process of RSPIs data collection by using the advanced technologies between the vehicle, infrastructure, driver and system of Internet of Vehicles (IoV) is possible in the near future. Recent progress in the development of Artificial Intelligence (AI) tools, supported by the development of cloud computing technologies and 5G mobile communication networks is a strong driving factor for upgrading traditional Vehicular Ad-hoc Networks (VANETs) into flexible heterogeneous IoV global communication architectures which are expected to satisfy strict communication requirements related to the networking of a wide range of entities (vehicles, pedestrians, infrastructure equipment, personal devices, sensors, etc.) for the needs of future IoV applications. This requires mass involvement of public decision makers, research and scientific community, vehicle manufacturers and suppliers, traffic information service providers, etc. Such a level of coordination and participation will be necessary for the provision of the pace and the critical mass of road safety data required for a comprehensive and in-depth analysis of road safety situation in a territory, detection of emergence problems at an earlier stage, evaluation of road safety measures, exchange of best knowledge, etc. Under this assumption, it is possible to generate large amounts of KPIs data, obtained from various projects, naturalistic driving studies, field operational tests, smart cameras, advanced smart solutions in urban area, smart in-car solutions, etc. This process will enable the management of safety performance indicators, especially management of driver's behavior in real-time, which can lead to a reduction of harmful impacts of traffic in the transition period. Finally, data-driven road safety management have a great potential for defining goal-oriented actions which contribute the better identification of the road safety problem, precisely defining countermeasures and implementation of the most effective measures.

## 6. Concluding remarks

Risky driving has been constantly researched due to the consequences it has on road safety. Risk of crashing increases as consequence of risky driving behaviors such as: drunk driving, over speeding, mobile phone use while driving, unfastened seat belts, overtaking on the wrong side, using two lanes, driving through an orange light that is turning red, talking on mobile phones, changing lanes without signaling, etc. Most of these factors can be measured and quantified using traditional and modern methods (unmanned aerial vehicle, data from automated vehicles and vehicle-to-everything communication-V2X, etc.), which enables the assessment of the driver's behavior as well as road safety level at the territory.

Star rating of safety performance indicators, especially indicators related to the driver's behavior, is a useful tool for monitoring road safety progress, defining earlier goal-oriented road safety actions for road safety improvement, identifying the critical factors leading to road accidents, cross-territory comparison and strengthening the proactive road safety management at the territory (both at the international and national level). Also, a star rating serves to decision-makers to help them recognize and understand the road safety-related problems and contributing to the practicality and resource-saving in the monitoring process.

The paper explores the star rating of the driver's behavior in order to provide credible road safety monitoring and identification of the factors that most contribute to risky driving behaviors at the 27 PAUs in the Republic of Serbia. The results are analyzed regarding the following aspects: 1) star rating of driver's behavior at the national level and 2) defining the factors monitoring phases to prevent risky driving for each PAU analyzed. The results have revealed that the star rating of driver's behavior is reliable and robust enough to enable defining preventive measures for road safety system improvement and national cross-territory comparisons, promptly. The most significant behavioral factors associated with safety outcomes that have been identified following this approach ensure the optimum choice of behavioral factors. In the case of the Republic of Serbia, we have identified that "alcohol", "mobile phone" and "protective systems-seat belts and child restraints are the key areas of concern and should be prioritized in road safety initiatives. Comparisons of results (comparison of PAUs' rankings) verify the robustness of the most significant factors and improve the credibility and interpretability of the star rating with a limited number of factors. The second part of the analysis has identified the most significant factors per PAU as well as defining the factors monitoring phases. The results presented show that there is a strong linear dependence between the overall star rating scores and values of the star rating with a limited number of factors and that the applied methodology for star rating driver's behavior provides a reliable and credible national cross-territory comparison. Furthermore, the results have shown that it is possible to identify the most contributed factors associated with risky driving behaviors for each police administration unit and define the stages of periodical monitoring in order to manage driver's behavior and to prevent risky driving. With a restricted road safety budget, the territories will be able to more easily decide on/opt for investing in actions that target the factors that most contribute to risky driving behaviors, with the goal of making more efficient expenditures in road safety improvement.

The process of upgrading the star rating methodology should be observed through the following limitations:

- Precision and reliability in selecting the factors that most contribute to risky driving behaviors increases proportionally with the number of factors involved (a more precise list of the factors that most contribute to risky driving behaviors is obtained on the basis of multiple factors analyzed);
- Disputable quality of available data, diversity in definitions of factors (esp. factors related to the alcohol), as well as a weak correlation dependence of input elements with final outputs. A more reliable analysis, with a higher degree of correlation with the final outputs can be expected when analyzing data over a longer period (e.g., 10 years) or when expanding the set of data obtained by automatic data collection;
- When it comes to the in-depth analysis of the driver's behavior at the national level, it is important to highlight the limitation concerning the value of factors for those time periods for which measurements have not been carried out in the field. Future research studies should tend to use those factors that are measured twice a year, unless otherwise indicated by the national monitoring methodology.

Further research should include star rating based on the other data sets i.e., factors related to the vulnerable road users. In that case, the overall star rating score and identification of the most significant factors will be obtained from the larger set of factors which ensures a more detailed analysis of contributory factors associated with risky driving behaviors. The development of the star rating for assessing road safety performance of a territory should be a possible game-changer for systematic management (data-driven road safety system) of road user's behavior, especially in case of automated process of RSPIs data collection.

**Declaration of Competing Interest**

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

**Data availability**

Data will be made available on request.

## Acknowledgments

The authors would like to thank the Editor-in-Chief, PhD Samuel G. Charlton, and several anonymous reviewers for their valuable comments and suggestions, which have been of great help in improving the quality of this paper. Also, the authors would like to thank the Road Traffic Safety Agency of the Republic of Serbia for providing data on safety performance indicators (factors associated with risky driving behavior).

## References

Aigner-Breuss, E. (2017). *Cost-benefit analysis – Seatbelts campaign, European road safety decision support system, developed by the H2020 project SafetyCube*. Retrieved from www.roadsafety-dss.eu on 26 July 2022.
Al-Haji, G. (2007). *Road safety development index (RSDI): Theory, philosophy and practice*. Department of Science and Technology. Linköping University, http://ir.nmu.org.ua/bitstream/handle/123456789/126470/0bb47fc915be24e29d6a9f7912a5abe3.pdf?sequence=1.
Babaee, S., Toloo, M., Hermans, E., & Shen, Y. (2021). A new approach for index construction: The case of the road user behavior index. *Computers & Industrial Engineering, 152*, Article 106993. https://doi.org/10.1016/j.cie.2020.106993
Babaee, S. (2022). *Developing a composite index for assessing safety performance of drivers*. PhD Thesis. Hasselt: School of Transportation Sciences, Hasselt University.
Bax, C., Wesemann, P., Goldenbeld, C., Wegman, F., Aarts, L., Gitelman, V., et al. (2012). *Developing a road safety index. Deliverable 4.9 of the EC FP7 project DaCoTA*. http://www.dacota-project.eu/Deliverables/DaCoTA_D4.9_developing%20a%20RSI%20deliverable.pdf.
Böhm, M., Franklin, J. R., Jones, S., Kovacikova, T., Nowicka, K., & Yannis, G. (2018), *Analysis of the state of the art, barriers, needs and opportunities for setting up a transport research cloud*. Brussels: European Commission Directorate-General for Research and Innovation. [Online] Available at: doi: 10.5281/zenodo.1469661.
Carson, J., Jost, G., Meinero, M. (2023). *Ranking EU progress on road safety: 17th Road Safety Performance Index report*. Brussels: European Transport Safety Council. [Online] Available at: https://etsc.eu/17th-annual-road-safety-performance-index-pin-report/.
Charnes, A., Cooper, W. W., Lewin, A. L., & Seiford, L. M. (1994). *DEA: Theory, Methodology, and Application*. Boston: Kluwer Academic Publishers.
Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision- making units. *European Journal of Operational Research, 2*, 429–444. http://www.utdallas.edu/~ryoung/phdseminar/CCR1978.pdf.
Chen, F., Shi, L., Li, Y., Wang, Q., Sun, H., Tang, X., et al. (2022). Measuring road safety achievement based on EWM-GRA-SVD: A decision-making support system for APEC countries. *Knowledge-Based Systems*. https://doi.org/10.1016/j.knosys.2022.109373
Chen, F., Wu, J., Chen, X., Wang, J., & Wang, D. (2016). Benchmarking road safety performance: Identifying a meaningful reference (best-in-class). *Accident Analysis and Prevention, 86*, 76–89. https://doi.org/10.1016/j.aap.2015.10.018
Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ, NJ: Lawrence Erlbaum.
Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155
Cooper, W. W., Seiford, L. M., & Zhu, J. (2004). *Handbook on data envelopment analysis*. Boston, MA: Kluwer Academic Publishers.
Daniels, S., Martensen, H., Schoeters, A., Van den Berghe, W., Papadimitriou, E., Ziakopoulos, A., et al. (2019). A systematic cost-benefit analysis of 29 road safety measures. *Accident Analysis & Prevention, 133*. https://doi.org/10.1016/j.aap.2019.105292
European Commission (2019). *EU road safety policy framework 2021–2030 - next steps towards "Vision Zero"*. Commission Staff Working Document SWD (2019) 283. Brussels: European Commission Directorate-General Mobility and Transport. https://ec.europa.eu/transport/road_safety/sites/default/files/move-2019-01178-01-00-en-tra-00_3.pdf.
Ghosal, A., & Conti, M. (2020). Security issues and challenges in V2X: A survey. *Computer Networks, 169*. https://doi.org/10.1016/j.comnet.2019.107093
Gitelman, V., Doveh, E., & Hakkert, S. (2010). Designing a composite indicator for road safety. *Safety Science, 48*, 1212–1224. https://doi.org/10.1016/j.ssci.2010.01.011
Haghani, M., Behnood, A., Dixit, V., & Oviedo-Trespalacios, O. (2022). Road safety research in the context of low-and middle-income countries: Macro-scale literature analyses, trends, knowledge gaps and challenges. *Safety Science, 146*, Article 105513. https://doi.org/10.1016/j.ssci.2021.105513
Haghani, M., Behnood, A., Oviedo-Trespalacios, O., & Bliemer, M. C. (2021). Structural anatomy and temporal trends of road accident research: Full-scope analyses of the field. *Journal of Safety Research, 79*, 173–198. https://doi.org/10.1016/j.jsr.2021.09.002
Haque, M. M., Oviedo-Trespalacios, O., Sharma, A., & Zheng, Z. (2021). Examining the driver-pedestrian interaction at pedestrian crossings in the connected environment: A Hazard-based duration modelling approach. *Transportation Research Part A: Policy and Practice, 150*, 33–48. https://doi.org/10.1016/j.tra.2021.05.014
Hermans, E. (2009). *A methodology for developing a composite road safety performance index for cross-country comparison*. University of Hasselt. PhD Thesis.
Hermans, E., Brijs, T., Wets, G. (2010). Bringing structure into road safety evaluation: A hierarchy of indicators. In *Proceedings of the 15th international conference road safety on four continents* (Vol. 785, pp. 273–281). Retrieved November 21, 2014, from University of Hasselt: https://doclib.uhasselt.be.
Hermans, E., Van den Bosshe, F., & Wets, G. (2008). Combining road safety information in a performance index. *Accident Analysis and Prevention, 40*, 1337–1344. https://doi.org/10.1016/j.aap.2008.02.004
Jameel, A. K., & Evdorides, H. (2021). Developing a safer road user behavior index. *IATSS Research, 45*(1), 70–78. https://doi.org/10.1016/j.iatssr.2020.06.006
Kaiser, S. (2017). *Cost-benefit analysis – Booster seat program, European Road Safety Decision Support System, developed by the H2020 project SafetyCube*. Retrieved from www.roadsafety-dss.eu on 26 July 2022.
Khakzar, M., Bond, A., Rakotonirainy, A., Trespalacios, O. O., & Dehkordi, S. G. (2021). Driver influence on vehicle trajectory prediction. *Accident Analysis & Prevention, 157*. https://doi.org/10.1016/j.aap.2021.106165
Koornstra, M., Lynam, D., Nillson, G., Noordzij, P., Petterson, H.-E., Wegman, F., et al. (2002). SUNflower: *A comparative study of the development of road safety in Sweden, the United Kingdom and the Netherlands.* Leidschendam: SWOV Institute for Road Safety Research. http://ec.europa.eu/transport/roadsafety_library/publications/sunflower_report.pdf.
Kopelias, P., Demiridi, E., Vogiatzis, K., Skabardonis, A., & Zafiropoulou, V. (2020). Connected & autonomous vehicles–environmental impacts – A review. *Science of the Total Environment, 712*(135237), 1–8. https://doi.org/10.1016/j.scitotenv.2019.135237
Kukić, D., Lipovac, K., Pešić, D., & Vujanić, M. (2013). Selection of a relevant indicator–road casualty risk based on final outcomes. *Safety Science, 51*(1), 165–177. https://doi.org/10.1016/j.ssci.2012.06.016
Meesmann, U., Wardenier, N., Torfs, K., Pires, C., Delannoy, S., Van den Berghe, W. (2022). *A global look at road safety. Synthesis from the ESRA2 survey in 48 countries. ESRA project (E-Survey of Road users' Attitudes)*. Brussel, Belgium: Vias Institute. [Online] Available at: https://www.esranet.eu/en/publications-2/esra2-publications/#main_reports.
Nardo, M., Saisana, M., Saltelli, A., & Tarantola, S. (2005). *Handbook on constructing composite indicators: Methodology and user guide*. Retrieved July 20, 2022, from OECD Statistics Working Papers 2005/03; OECD Publishing. doi: 10.1787/533411815016.
Nguyen-Phuoc, D. Q., Oviedo-Trespalacios, O., Vo, N. D. Q., Le, Q. T., Su, D. N., & Cao, M. X. (2020). Prevalence and factors associated with turn signal neglect-related crashes among motorcyclists and car drivers in Vietnam. *Transportation Research Part F: Traffic Psychology and Behaviour, 73*, 38–49. https://doi.org/10.1016/j.trf.2020.06.003
OECD/ITF (2020). Road safety annual report 2020. Paris. https://www.itf-oecd.org/road-safety-annual-report-2020.
Outay, F., Mengash, H. A., & Adnan, M. (2020). Applications of unmanned aerial vehicle (UAV) in road safety, traffic and highway infrastructure management: Recent advances and challenges. *Transportation Research Part A: Policy and Practice, 141*, 116–129. https://doi.org/10.1016/j.tra.2020.09.018
Oviedo-Trespalacios, O., & Scott-Parker, B. (2018). Young drivers and their cars: Safe and sound or the perfect storm? *Accident Analysis & Prevention, 110*, 18–28. https://doi.org/10.1016/j.aap.2017.09.008

Pallant, J. (2011). *SPSS survival manual*. McGraw-Hill Education (UK). http://eunacal.org/metodakerkimi/wp-content/uploads/spss/SPSS_Survival_Manual_4th_Edition.pdf.

Papadimitriou, E., & Yannis, G. (2013). Is road safety management linked to road safety performance? *Accident Analysis and Prevention, 59*, 593–603. https://doi.org/10.1016/j.aap.2013.07.015

Permanent International Association of Road Congresses. (2003). Road Safety Manual: Recommendations from the World Road Association (PIARC). Route2 Market.

Pešić, D. R. (2012). Developing and improving the method for measuring the level of road safety at the territory. PhD Thesis. Belgrade: Faculty of Transport and Traffic Engineering.

Pires, C., Torfs, K., Areal, A., Goldenbeld, C., Vanlaar, W., Granié, M. A., et al. (2020). Car drivers' road safety performance: A benchmark across 32 countries. *IATSS Research, 44*(3), 166–179. https://doi.org/10.1016/j.iatssr.2020.08.002

Rahman, H. M., & Abdel-Aty, M. (2021). Application of connected and automated vehicles in a large-scale network by considering vehicle-to-vehicle and vehicle-to-infrastructure technology. *Transportation Research Record, 2675*(1), 93–113. https://doi.org/10.1177/0361198120963105

Scott-Parker, B., & Oviedo-Trespalacios, O. (2017). Young driver risky behavior and predictors of crash risk in Australia, New Zealand, and Colombia: Same but different? *Accident Analysis & Prevention, 99*, 30–38. https://doi.org/10.1016/j.aap.2016.11.001

Shen, Y., Hermans, E., Ruan, D., Wets, G., Brijs, T., & Vanhoof, K. (2011). Modeling qualitative data in data envelopment analysis for composite indicators. *International Journal of System Assurance Engineering and Management, 2*(1), 21–30. 10.1007/s13198-011-0051-z#page-1.

Sheykhfard, A., & Haghighi, F. (2019). Performance analysis of urban drivers encountering pedestrian. *Transportation Research Part F: Traffic Psychology and Behaviour, 62*, 160–174. https://doi.org/10.1016/j.trf.2018.12.019

Sheykhfard, A., & Haghighi, F. (2020). Driver distraction by digital billboards? Structural equation modeling based on naturalistic driving study data: A case study of Iran. *Journal of Safety Research, 72*, 1–8. https://doi.org/10.1016/j.jsr.2019.11.002

Tešić, M. (2018). *Road safety assessment based on a road safety performance index*. Belgrade: Faculty of Transport and Traffic Engineering. PhD Thesis.

Tešić, M., Hermans, E., Lipovac, K., & Pešić, D. (2018). Identifying the most significant indicators of the total road safety performance index. *Accident Analysis and Prevention, 113*, 263–278. https://doi.org/10.1016/j.aap.2018.02.003

Watson, C. E., & Austin, R. A. (2021). Differences in rural and urban drivers' attitudes and beliefs about seat belts. *Accident Analysis & Prevention, 151*. https://doi.org/10.1016/j.aap.2021.105976

Wegman, F., Commandeur, J., Doveh, E., Eksler, V., Gitelman, V., Hakkert, S., et al. (2008). *SUNflowerNext: Towards a composite road safety performance index*. Leidschendam: SWOV Institute for Road Safety Research.

Wegman, F., Eksler, V., Hayes, S., Lynam, D., Morsink, P., & Oppe, S. (2005). *SUNflower+6. A comparative study of the development of road safety in the SUNflower+6 countries: Final report*. Leidschendam: SWOV Institute for Road Safety Research.

Wegman, F., & Oppe, S. (2010). Benchmarking road safety performances of countries. *Safety Science, 48*, 1203–1211. https://doi.org/10.1016/j.ssci.2010.02.003

World Health Organization. (2018). Global status report on road safety 2018: Summary (No. WHO/NMH/NVI/18.20). World Health Organization. [Online] Available at: https://www.who.int/publications/i/item/9789241565684.

World Health Organization. (2022). *Global plan for the decade of action for road safety 2021–2030 (No. WPR/2022/VNM/001)*. WHO Regional Office for the Western Pacific. [online] Available at: https://www.who.int/teams/social-determinants-of-health/safety-and-mobility/decade-of-action-for-road-safety-2021-2030.

Yager, R. R. (1988). On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Transactions on Systems, Man and Cybernetics, 18*, 183–190.

Yannis, G., Folla, K., Nikolaou, D., Dragomanovitis, A., & Wang (2020). Development of a platform for global road safety data analysis. In *Proceedings of the 8th transport research arena TRA 2020 conference (Helsinki, Finland, 27–30 April 2020)*. Available at: https://www.nrso.ntua.gr/geyannis/pub/pc369-development-of-a-platform-for-global-road-safety-data-analysis/.