ISSN 1648-4142 / eISSN 1648-3480

TRANSPORT

2025
Volume 40
Issue 4
Pages 327–342
https://doi.org/10.3846/transport.2025.20539

Original Article

NOVEL HYBRID MODEL FOR ADDRESSING UNCERTAINTY OF THE ROAD SAFETY COMPOSITE INDICATOR: INTEGRATION OF DEA AND WEIGHTED GRA

Boris ANTIĆ[1], Mirjana GRDINIĆ-RAKONJAC[2]✉, Vladimir PAJKOVIĆ[2]

[1] *Faculty of Transport and Traffic Engineering, University of Belgrade, Belgrade, Serbia*
[2] *Faculty of Mechanical Engineering, University of Montenegro, Podgorica, Montenegro*

Highlights:

- DEA is widely used method for various studies considering traffic safety, especially for constructing one overall composite indicator that will include all available information on road safety and represent safety performance;
- weighted GRA is used as extension to the CE method for creating a composite index of territories and for further differentiation of obtained scores to a full and rational ranking of entities;
- hybrid DGF methodology provides an enriched picture compared to the basic DEA approach, softens the discrimination of inputs and outputs usually conducted by DEA and enables decision-makers to reduce self-evaluation, including peers in assessment;
- DGF can be used when the number of inputs and outputs (road safety indicators) is higher than the number of entities (territories under evaluation) because the final results are not affected by the number of DMUs;
- DGF can be considered as a natural extension of DEA and CE, and the verified robustness of the results indicates the feasibility of solving multi-attribute performance evaluation problems in many other fields besides road safety.

Article History:

- submitted 16 February 2021;
- resubmitted 18 May 2021,
- accepted 14 August 2021.

Abstract. Creating a composite indicator is a popular concept in evaluating and comparing road safety of territories. While many other approaches are used, most current interest is on Data Envelopment Analysis (DEA) for measuring the relative performance in efficiency terms. Usefulness of DEA, which measure performance in efficiency terms, has already been proven. However, the indicators commonly used to construct a road safety composite index are not always precise and accurate (particularly safety performance indicators), and the results obtained from this type of data does not seem to be solid. So, the main aim of this paper is to represent the novel methodology to effectively evaluate the state of road safety and create a reliable composite index of the selected entities when imprecise data are involved and, to produce reliable composite index under uncertain environment. DEA, the weighting method Fan–Ma, and the Grey Relational Analysis (GRA) were integrated into hybrid methodology to obtain a more realistic and relevant picture of road safety. Applying this hybrid methodology, peculiarity of DEA is retained, scores are further differentiated, and entities are ranked and classified according to the road safety level. A case study was conducted to evaluate and rank municipalities, determine road safety classes and benchmark the territories under study. Results were verified indicating the robustness and effectiveness of the proposed methodology and its superiority to basic DEA as regards the territory ranking.

Keywords: road safety, composite indicator, data envelopment analysis, uncertainty, grey relational analysis, Fan–Ma.

✉ Corresponding author. E-mail: *grdinicm@ucg.ac.me*

Notations

B/W – best–worst;
CE – cross efficiency;
DEA – data envelopment analysis;
DGF – DEA-grey-fan-ma safety score;
DMU – decision-making unit;
Fan–Ma – weighting method developed by Z.-P. Fan, and J. Ma;
G/E – grey entropy;
GRA – grey relational analysis;
PCA – principal component analysis;
MCDM – multi-criteria decision-making;
wDEA – weighted DEA.

1. Introduction

Comparing road safety performance among territories (municipalities, regions, countries) and learning from better ones has recently been of great interest to experts, since it has been confirmed to be a useful tool for policy makers. Finding a fair and reliable approach to represent safety performance on territory, the most often implies constructing an overall composite indicator that will include all available information. Selecting road safety indicators as input variables, imputing missing data (if necessary), normalizing data, assigning weights and aggregation were listed as main steps for constructing a composite index (OECD 2008). Usually, traditional indicators used for describing state of road safety relies on precise data such as number of fatalities and injured, number of accidents, motorization level, etc. In contrast, safety performance indicators that describe drivers' attitudes and behaviour (Bastos *et al.* 2015; Jameel *et al.* 2021; Babaee *et al.* 2021) are usually collected by measuring with physical devices and from questionnaires. Since not every driver's behaviour has been recorded, these data cannot be considered fully accurate. Furthermore, qualitative information collected by questionnaires cannot be fully described with numerical data and it is impossible to take them as completely accurate (Zhu, Hipel 2012). Since uncertainty is the inherent feature of this mentioned kind of data, ignoring it in the performance evaluation will lead to unreliable results; therefore, it is needed to consider uncertain parameters in the process of constructing reliable road safety composite index.

Combining road safety data into one index can be done by various methods. Although the application of MCDM is particularly popular, there is no single standardized and recommended method. Hermans *et al.* (2008) tested different MCDM techniques for this purpose and prioritized DEA among them. Since then, DEA has been most commonly used to evaluate relative road safety performance in efficiency terms conducting weighting and aggregation at the same time.

DEA measures the efficiency of entities, the so-called DMUs, based on weights that have not been a priory known. Each DMU in DEA uses its own combination of input and output weights (optimal weights set) to maximize its efficiency value. A frontier is formed by the efficient DMUs that represent the best practice. How entities in DEA will be classified depends on location relative to an efficient frontier. Inefficient DMUs can be compared with efficient ones that can be observed as benchmarks for further improvement. Although it has proven to be extremely useful technique, several drawbacks occur while using DEA. Firstly, number of DMUs under consideration are relative to the number of input and output and they should be in balance. In order to obtain usable results with DEA, Alper *et al.* (2015) highlights that the number of DMUs should be (minimum) 3 times greater than total number of inputs and outputs. Nevertheless, most of the time a few DMUs will achieve score of one (efficient DMUs) and prevents final ranking among them.

In addition, classical DEA models take into account only fully known and precise input and output data (Peykani *et al.* 2019), which is not the case in real-world applications. The usefulness of DEA and reliability of calculated efficiency scores are reduced when imprecise data in matter and it is of great significance to preserve its robustness. Many authors have suggested different methods for handling imprecise information; however, there are no papers in the field of road safety evaluation of the territory addressing the uncertainty of DEA scores obtained from such data. Javed *et al.* (2019) stated that grey system theory and its models, especially GRA, are superiority to handle uncertainty for evaluation of a system when the decision-makers are not certain about the sufficiency and completeness of data at hand. To the best of our knowledge, no studies have been devoted to road safety assessment under uncertainty based on the DEA model features and also the GRA method, simultaneously, under the uncertain environment. So, the main aim of this study is to propose novel hybrid method to process road safety index score constructed from easily available data and obtained with DEA; and to construct reliable composite index for each territory under study. Such method is needed because most of the developing countries still do not have reliable road safety databases and also, do not have plenty of accurate traffic safety data. In other words, the main goal of this research was to fil that gap in the literature by creating the model usable in countries with modest data related to traffic accidents and violations in traffic.

The existing models for road safety evaluation commonly consider the uncertainty of data when subjective opinion of road safety experts is involved. However, this proposed approach accesses the unreliability of DEA results obtained with crisp data with questionable certainty and accuracy, providing a set of robust road safety indexes for each territory. The contributions of this paper also include: practical usefulness since it does not have complex algorithm and it is convenient and easy-to-solve without requiring anything more than excels' add-in solver; fully ranking of all territories under analysis without implementing subjective information from outside; handling the uncertainty of road safety DEA scores when imprecise, easy available data are involved; recognizing the best practice among territories; and applying proposed method in many other fields beside road safety, which has been verified by robust and logical results.

The rest of the paper is as follows:

- current Section 1 – an introduction;
- the next Section 2 provides a literature review;
- an overview of the applied methodologies is briefly presented in the Section 3, which ends with the identification and presentation of the road safety indicators for a case study;
- in the Sections 4, the results of the implemented methodology and benchmarking according to them are given;
- in the section 5 the results and findings of the proposed method are discussed;
- finally, Section 6 concludes the paper and represents the direction of future research.

2. Literature review

DEA is integral part of MCDM methods and has been widely used in many fields. The versatility of DEA is best illustrated by a number of studies that have attempted to provide a comprehensive review of its' application in a particular field, such Dutta *et al.* (2022) in supplier selection, Tsaples *et al.* (2021) in sustainability, Rostamzadeh *et al.* (2021) in benchmarking, Mahmoudi *et al.* (2020) and Mardani *et al.* (2016) in transportation, Zakowska & Godycki-Cwirko (2020) in health care system, Mardani *et al.* (2017) in energy efficiency, and others. Considering the field of road traffic and transportation, DEA has been used for evaluating the performance of public transport organizations (Geerlings *et al.* 2006, Yushimito *et al.* 2018). addressing the efficiency of transport sector (Baležentis, A., Baležentis, T. 2011), evaluating the performance of older drivers (Babaee *et al.* 2015), for evaluation of risky highway segments and contributing factors (Raheel Shah *et al.* 2019) and, in particular, for constructing an overall performance index score to evaluate the safety level.

Hermans *et al.* (2008) was one of the 1st authors who evaluated safety performance of territories with composite index obtained with DEA. DEA was applied to analyse the performance of territory on national level (Shen *et al.* 2013; Bastos *et al.* 2015; Nikolaou, Dimitriou 2018), micro level (Antić *et al.* 2020; Kang, Wu 2022; Folla *et al.* 2021) and also for identifying roads where the needs to improve safety are the greatest (Fancello *et al.* 2020). Because of the in-consistency of chosen indicators and the dependence of the results on them, Shen *et al.* (2012) and Rosić *et al.* (2017) introduced extensions of the DEA model in order to select the most optimal method for creating a composite index. Further, Shen *et al.* (2011) and Babaee *et al.* (2021) used a multi-layer DEA model to integrate a hierarchical structure of input and output indicators.

Rassafi *et al.* (2018) claimed that every safety analysis has some degree of uncertainty. So, many authors in the literature are finding ways to overcome these issues resulting in a number of methods such as robust optimization (Mirzapour Al-E-Hashem *et al.* 2011), rough sets (Xiong *et al.* 2018), entropy (Chen *et al.* 2015; Castro-Nuño, Arévalo-Quijada 2018), evidential reasoning approach (Ayati *et al.* 2012; Rassafi *et al.* 2018) widely used fuzzy set theory (Ma *et al.* 2011; Bao *et al.* 2012, Qazvini *et al.* 2016; Memiş *et al.* 2020; Đalić *et al.* 2021) and others, combining different MCDM methods (Nenadić 2019; Stanković *et al.* 2020).

When handling with data that are generally imprecise, that often involves uncertainties of different kinds, including ignorance, fuzziness, inaccuracy or vagueness. In most transportation studies, observed uncertainty was due to the involvement of subjective opinion of decision-makers and it was successfully resolved using fuzzy theory. However, some surveys have revealed that fuzzy sets are not always a suitable tool for modelling imprecise data (Liu 2012, Jiang *et al.* 2019) such as, the most commonly, easily available road safety data. In addition, the GRA, as a part of grey theory, was suggested to deal with uncertainty of imprecise and vague information in many fields. Yang *et al.* (2019) states that grey systems are an ideal option in capturing objective uncertainty like information incompleteness and inaccuracy not only for small data but also big data. Considering road safety, Lu & Wevers (2007) used GRA to rate scenarios that will be implemented to improve road safety in urban roads. Hu *et al.* (2010) used GRA in addition to fuzzy and DEA to develop a public transport network evaluation. René *et al.* (2016) made predictions regarding the expected number of accidents using a time series and GRA, and, based on the results, suggested target improvements and measures to be conducted. Both, Liu *et al.* (2017) and Grdinić-Rakonjac *et al.* (2021) used GRA to calculate weights of road safety indicators for purpose of combining them in one single safety score.

Application of DEA as methodology for constructing a composite index of the territory, gives the researcher an opportunity to handle uncertainties in 3 different points: input and output data (Behnood *et al.* 2017; Ehrgott *et al.* 2018; Amini *et al.* 2019); function constraints (Shen *et al.* 2015; Blagojević *et al.* 2020; Mitrović Simić *et al.* 2020; Babaee *et al.* 2021) and results (CE matrix). Since Sexton *et al.* (1986) developed a CE as an extension to DEA to provide an ordering among DMUs and to eliminate unrealistic weight allocation without applying subjective restrictions from outside (e.g., from experts), it has gained interest of many authors. CE includes peer evaluations of DMUs rather than pure self-evaluation and for accessing uncertainty of these scores, authors usually use statistics of centralization (mean value). Anderson *et al.* (2002) found that this method can be unrealistic as well which resulted in many different CE extensions that were suggested based on Shapley value and Ordered Weighted Averaging operators (Wang, Chin 2011) Shannon entropy (Lee 2019; Su, Lu 2019), Bonferroni mean (Behdani, Darehmiraki 2022), GRA (Si, Ma 2019), and others. GRA is common method used to address data inaccuracy and its aggregation with DEA can be found in the literature: Markabi & Sabbagh (2014), Pakkar (2016), Küçükönder *et al.* (2019), Tsolas (2019), etc. Since independence of the GRA from the large sample size has already been proven and its ability to handle data with uncertain sufficiency and completeness has been confirmed, in this paper we propose its' implementation at the 3rd DEA point, the CE matrix. In addition, for purpose to evaluate state of road safety and to create composite index of the territory, DEA CE extension based on the weighted GRA, is suggested. The main aim of this hybrid model is to cope with the uncertainty of the results derived from imprecise, easily available road safety data, particularly with the uncertainty of scores obtained with DEA.

Given that the final index of each territory must resolve incorporated uncertainty inherent in the aggregation of CE, in this paper GRA is applied to combine the entire range of performance efficiencies in CE matrix into one, single value. Relevant researches show that GRA can directly evaluate peer efficiencies instead of making an assumption for original distributions of data, which indicates that GRA is a robust and convenient (Wang, Yao 2018). With this novel methodology, by processing the uncertainty of the results, it is possible to create a reliable composite safety index of the territory and additionally establish a more rational ranking among the territories under evaluation. To illustrate applicability of proposed hybrid methodology, it is conducted to evaluate road safety of 21 territories in Montenegro.

3. Methods

The research flow with the created methodology for constructing reliable road safety composite index from imprecise data is shown in Figure 1. Each step is conditioned by the previous one and represents the input parameter for the next step of the model.

3.1. DEA

DEA is a widespread non-parametric linear programming technique for measuring relative efficiencies of DMUs operating in multiple input and output environment. According to the standard DEA methodology (developed by Charnes *et al.* (1978) and extended by Banker *et al.* (1984), the efficiency score is defined as follows: if m-dimensional inputs and s-dimensional outputs are included for each

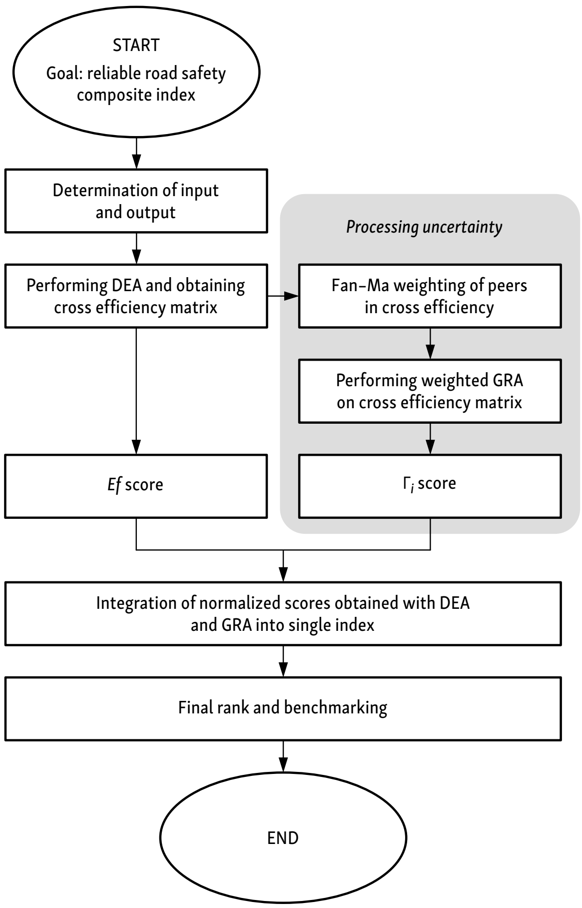

Figure 1. Framework of the proposed methodology

$DMU_j \left(j = 1, \ldots, n\right)$, the mathematical formulation in linear form of the basic input- oriented DEA model is given as:

$$\max Ef_p = \sum_{r=1}^{s} u_r \cdot y_{rp}, \tag{1}$$

subject to:

$$\sum_{i=1}^{m} v_i \cdot x_{ip} = 1;$$
$$\sum_{r=1}^{s} u_r \cdot y_{rj} \le \sum_{i=1}^{m} v_i \cdot x_{ij},\ \ j = 1, \ldots, n;$$
$$u_r \ge 0,\ \ r = 1, \ldots, s;$$
$$v_i \ge 0,\ \ i = 1, \ldots, m,$$

where: y_{rj} denotes the rth output and x_{ij} denotes the ith input of the jth DMU; u_r is the weight of the output r; v_i is the weight for the input i.

It is important to mention that considering road safety, outputs should be as low as possible, and in order to address this issue, input and output data in the DEA model for calculating composite road safety index are reciprocal values. When a set of optimal weights $\left(u_r^*, v_i^*\right)$ for DMU_p $\left(p = 1, \ldots, n\right)$ under consideration is calculated by Model (1), then CE of DMU_j relative to DMU_p (peer evaluation), is:

$$Ef_{jp} = \frac{\sum_{r=1}^{s} u_r^* \cdot y_{rj}}{\sum_{i=1}^{m} v_i^* \cdot x_{ij}}. \tag{2}$$

In that way, a $n \times n$ matrix of efficiency is constructed where basic DEA efficiency Ef_p (self-evaluated) is placed in the matrix diagonal. When the CE matrix is constructed, the next step, in this paper, is calculating the grey relational degree that represents the similarity measure between the efficiencies of DMUs with peers' weights and an ideal reference set (the set of the highest efficiencies). Detailed steps of this procedure are presented in the next section.

3.2. GRA

Grey theory, firstly proposed by Deng (1989), is a useful model for the analysis of uncertain systems with partially known and partially unknown information when multiple attributes matter. GRA, as a part of a grey system theory, solves multi-attribute decision-making problems by aggregating multiple attribute values into a single value for each alternative. It is a normalization-based technique, which implies positive values of the data sequence generated by translating the performance of all alternatives (in our case, DEA CE scores of all DMUs) into series.

If we have n data sequences and m attributes, and assuming positive criteria, normalization of decision matrix $X_j = x_j(1),\ x_j(2),\ x_j(k),\ \ldots,\ x_j(m)$, where: $x_j(k)$ is the efficiency score of DMU_j $\left(j = 1, \ldots, n\right)$ relative to DMU_k $\left(k = 1, \ldots, l\right)$ is conducted as follows:

$$g_j(k)=\frac{x_j(k)\cdot\min_{j=1}^{n}x_j(k)}{\max_{j=1}^{n}x_j(k)\cdot\min_{j=1}^{n}x_j(k)}. \quad (3)$$

Accordingly, the normalized matrix is derived as G:

$$G=\begin{bmatrix} g_1(1) & g_1(2) & \dots & g_1(l) \\ g_2(1) & g_2(2) & \dots & g_2(l) \\ \dots & \dots & \dots & \dots \\ g_n(1) & g_n(2) & \dots & g_n(l) \end{bmatrix}. \quad (4)$$

Here, the G matrix is a normalized CE matrix derived by DEA. Peer DMUs are in rows and target DMUs that are maximized are in columns. Next, the referential set representing the virtual ideal set and constructed from ideal values of attributes from the normalized matrix is generated as $G_0=(g_{01},g_{02},\dots,g_{0l})$. The referent set selected from the CE matrix represents the highest efficiency of every DMU, which is basically a value at the diagonal of the normalised matrix. Further, the similarity of every CE score relative to the ideal set is calculated as the grey relational coefficient as follows:

$$\zeta_{0j}(k)=\frac{\Delta_{\min}+\rho\cdot\Delta_{\max}}{\Delta_{0j}(k)+\rho\cdot\Delta_{\max}}, \quad (5)$$

where:

$$\Delta_{0j}(k)=\left|g_0(k)-g_j(k)\right|;$$

$$\Delta_{\max}=\max_{j=1}^{n}\max_{k=1}^{l}\Delta_{0j}(k);$$

$$\Delta_{\min}=\min_{j=1}^{n}\min_{k=1}^{l}\Delta_{0j};$$

$$\rho=[0,1]$$

with ρ representing the distinguish coefficient.

To indicate a moderate distinguish effect, ρ has been assigned a value of 0.5. In the next step, the grey relational degree is derived as the average of grey relational coefficients with Equation (6) or as a weighted sum with Equation (7) if weights $w(k)$ for attributes are known:

$$\Gamma_{0j}(k)=\frac{1}{l}\cdot\sum_{k=1}^{l}\zeta_{0j}(k); \quad (6)$$

$$\Gamma_{0j}(k)=\frac{1}{l}\cdot\sum_{k=1}^{l}\zeta_{0j}(k)\cdot w(k). \quad (7)$$

In order to further differentiate the influence on final road safety composite index and determine the priority for the set of the weights obtained by DEA with self-evaluation of each municipality, calculating weighted grey relational degree is suggested. For this purpose, in the next section, steps of the Fan–Ma method are shown. This particular weighting method is chosen because it gives more accurate results for a smaller (up to 30) number of entities (Ma *et al.* 1999), which is mostly the case in creating a road safety composite index of the territory.

3.3. Fan–Ma weights

The well-known objective weighting method Fan–Ma, proposed by Fan *et al.* (1996) and described by Ma *et al.* (1999), is used for weight optimization for calculating the grey relational degree. An advantage of Fan–Ma is the higher accuracy of the results with a small data sample, which is the case here. The criterion weights are determined by optimization, as they pre-normalize the alternative ratings, and the decision matrix is transformed into a new weighted matrix $y=\left[y_{jk}\right]_{nxl}$, where $y_{jk}=w_k\cdot x_{jk}$; $j=1,\dots,n$; $k=1,\dots,l$.

After identifying the "ideal points" in the decision matrix as $A^*=(y_1^*,y_2^*,\dots,y_l^*)$, where $y_k^*=w_k\cdot x_k^*$; $x_k^*=\max(x_{1k},x_{2k},\dots,x_{nk})$, calculation of the distance from the ideal set can be defined as follows:

$$g_j=\sum_{k=1}^{l}\left(y_k^*-y_{jk}\right)^2=\sum_{k=1}^{l}\left(x_k^*-x_{jk}\right)^2. \quad (8)$$

Weight determination w_k is:

$$\min\sum_{j=1}^{n}g_j=w^T\cdot H\cdot w=1 \quad (9)$$

subject to:

$$e^T\cdot w=1,\ w\geq 0,$$

where: $w=(w_1,w_2,\dots,w_l)^T$; $e=(1,1,\dots,1)^T$; H represents the diagonal matrix with elements $h_{kk}=\sum\left(x_k^*-x_{jk}\right)^2$.

When Langragian $L=w^T\cdot H\cdot w+2\cdot\gamma\cdot\left(e^T\cdot w-1\right)$ is introduced and differentiated, the following equations are obtained:

$$H\cdot w+\gamma\cdot e=0;$$
$$e^T\cdot w=1 \quad (10)$$

with solutions:

$$w^*=\frac{H^{-1}\cdot e}{e^T\cdot H^{-1}\cdot e};$$
$$\gamma^*=\frac{-1}{e^T\cdot H^{-1}\cdot e}. \quad (11)$$

In this way, the optimization is simplified and effectively completed with the final calculation of attribute weights k obtained as:

$$w_k^*=\frac{1}{\left(\sum_{j=1}^{n}\left(x_k^*-x_{jk}\right)^2\right)\cdot\left(\sum_{k=1}^{l}\frac{1}{\sum_{j=1}^{n}\left(x_k^*-x_{jk}\right)^2}\right)}. \quad (12)$$

It must be noted that the sum of the final weights are equal to 1 and the obtained weights w_k are further normalized with:

$$W_k^*=\frac{w_k^*}{\sum w_k^*}. \quad (13)$$

After the calculation of Fan–Ma weights, the weighted grey correlation degree $\Gamma_j(k)$ is derived. In the following section, the procedure for obtaining the final road safety score (DGF) with proposed methodology is described.

3.4. Obtaining the final score

As given in scheme in Figure 1, after obtaining the DEA efficiencies, weights for all attributes (peer DMUs) are derived by Fan–Ma. Afterwards, the weighted grey relation degree is obtained (incorporating those weights) constructing correlation degree scores $\Gamma_j\left(k\right)$. Further, the final score by the proposed methodology is derived as a combination of EF_p (self-evaluated DEA score) and Γ_j (grey CE score), which is previously normalized with:

$$Ef_j^{\,+} = \frac{Ef_j}{\max\limits_{j=1}^{n} Ef_j};$$

$$\Gamma_j^{+} = \frac{\Gamma_j}{\max\limits_{j=1}^{n} \Gamma_j}. \tag{14}$$

Final efficiencies are given with the following equation:

$$DGF_j = \alpha \cdot Ef_j^{\,+} + \beta \cdot \Gamma_j^{+} \tag{15}$$

with $\alpha + \beta = 1$, where the preference degree of road safety experts is included with α and β, reflecting the preference to DEA scores and scores based on weighted GRA, respectively.

The final efficiencies are within the range 0 to 1 due to the normalization step by Equation (14). The ranks of all DMUs are determined, and, by taking reciprocal values as input and output, data ranking is performed in a way whereby entities with a higher score occupy higher rank positions. A main advantage of the proposed method is addressing uncertainty of scores obtained by DEA, which makes efficient DMUs and these with equal efficiencies to be further distinguished, ending with a more rational ranking of all DMUs in sense of self-evaluation and peer-evaluation and at the same time retaining peculiarity of DEA. Furthermore, this easy-to-solve methodology softens the discrimination of inputs and outputs usually conducted by DEA and enables decision-makers to reduce self-evaluation, including peers in assessment. In addition, it can be used when the numbers of inputs and outputs in DEA are larger than the number of DMUs, which is occasionally the case in evaluating the road safety of a territory. Efficient DMUs in DEA represent a benchmark for others underperforming DMUs for further analysis. Likewise, applying the proposed methodology, many efficient entities become inefficient. In order to define the benchmark entities (in our case, territories), grouping and comparison within a specific group are conducted as described by Chen *et al.* (2016).

3.5. Dataset for the case study

Describing the state of the territory with a composite index, the complexity and multidisciplinary context of road safety can be captured in single comprehensive measure (Gitelman *et al.* 2010). In this study, the proposed DGF method is implemented to evaluate road safety of municipalities in Montenegro. Since the final assessment mostly depends on the choice of indicators that will be aggregated (Hermans *et al.* 2008), selection of relevant ones is an important task. Authors usually make an effort to respect indicators' criteria, established by Hermans (2009), namely: understandable, measurable, relevant, data available, reliable, comparable, specific and sensitive. However, using the best indicators is limited due to the unavailability of comparable road safety data for different territories, so many authors use the best-available indicators. 3 domains that affect the occurrence and consequences of traffic accidents and were used as traditional approach (ETSC 2001) are traffic participants (behaviour), road infrastructure (environment) and vehicles. Within the 1st SafetyNET project (Koornstra *et al.* 2002), categorization was performed on 4 domains concerning the policy performance, final and intermediate outcomes and background characteristics of the territory. In his work, Al-Haji (2007) used 8 different indicators classified into safer system, safer people and safer product domain. In addition, Hermans *et al.* (2008) used a set of safety performance indicators developed by SafetyNET project and created a final comprehensive list of safety risk domains that mostly refers to the behaviour of traffic participants. Hermans' (2009) crucial categorisation on alcohol and drugs, speed, protective systems, visibility, vehicle, infrastructure and trauma management is one of the most commonly used safety performance indicators when creating a composite road safety index.

Relying to the literature, authors of this paper chose behaviour related indicators, speeding, driving under the influence of alcohol or drugs, driving while using a telephone and not using a seat-belt as input indicators in proposed DEA–GRA model. In addition, as final outcomes, number of accidents and number of fatalities were used as output indicators. To evaluate road safety of 21 municipalities and to illustrate applicability, the DGF method is conducted. Given that proposed methodology requires that all territories involved in the analysis need to be described with the same indicators and given the lack of availability of road safety data in mentioned county, number of traffic violations recorded by the police is considered as the best-available indicators that will describe road users' behaviour (Antić *et al.* 2020). It should be mentioned that some of the influential system indicators, such as roads conditions, trauma management, age of the vehicle fleet, etc., were not included in the analysis because they were unavailable for all territories under evaluation.

The final list of selected indicators available for every territory is presented in Figure 2. The data source is Ministry of the Interior (Montenegro) database and for purpose to capture real situation in Montenegro roads, a 4-year average number of all indicators is used (2011–2014). Some basic characteristics of selected country are given as follows (all data are collected from the Ministry of the Interior). Montenegro is a South-Eastern European country with a high motorization rate (330 cars per 1000 inhabitants) and a road density of 57.2 km/km^2 (in 2016). Montenegro recorded a 50% decrease in road deaths in 2018 compared to 2010, with the highest number of accidents in the capital Podgorica (PG) about 35% of the states' accidents occur.

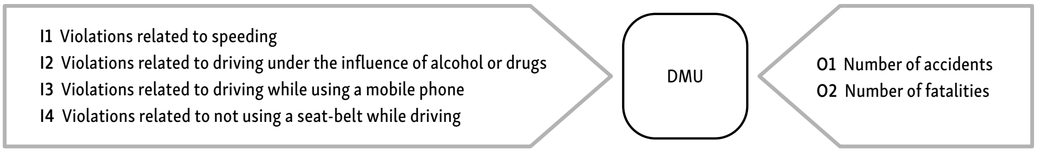

Figure 2. Selected inputs and outputs for the case study

The main objective of the DEA model is to calculate weights to make outputs as high as possible in order to maximize efficiency. However, it is very important to emphasize that considering road safety, outputs should be as low as possible, and because of that fact, input and output data in the proposed DEA–GRA model are reciprocal values. In addition, higher rank positions occupy municipalities with better performance. Moreover, although earlier mentioned Alpers' rule has been overcome with the proposed methodology (which will be shown later in the paper), road safety indicators were selected to satisfy that rule resulting with 6 indicators versus 21 analysed territories (Table 1).

4. Results

4.1. Ranking

Relative efficiencies for each municipality, related to the set of input and output data, were determined, and a CE matrix is constructed (Table A1 in Appendix A). Diagonal values of the CE matrix represent DEA self-evaluation scores, which indicated that almost half of the municipalities are efficient. This is due to the DEA flexibility in choosing weights for purpose to maximize the final efficiency score. In addition, zero weights are commonly assigned to several inputs and outputs so many DMUs appeared to be efficient, hence results cannot be reliable. Furthermore, DEA has failed to distinguish the territories. Calculating CE score by averaging peer efficiencies, the uncertainty of results was not considered. To further analyse the CE matrix in uncertain environment, the proposed DGF method based on weighted GRA was conducted.

Following the steps described in Section 3, the CE matrix was normalized, and Fan–Ma weights were derived (Table 2) in order to determine the priority for the set of the weights obtained by DEA with self-evaluation of each municipality. The obtained Fan–Ma weights indicate that the DG and BA weights sets have the greatest influence on final aggregated score, with an assigned value 0.059 and the lowest assigned weights of 0.028 have a KL weight set. Incorporating this information for calculating the final efficiency, the DEA peculiarity is retained and with further steps, uncertainty is addressed. In addition, the grey relational matrix was constructed (Table A2 in Appendix A), measuring closeness to an ideal reference set. An assumption was made that each municipality tended to be efficient, and the ideal reference set was therefore made of efficiencies equal to 1.

Further, Fan–Ma weights were integrated with grey relational coefficients, and the weighted efficiencies of each municipality $\left(\Gamma_i\right)$ were obtained (Table A3 in Appendix A). As next step, the grey relational degree is conducted by averaging weighted coefficients. Grey relational degree presents peer-included evaluation score $\left(\Gamma_i\right)$ of each territory taking in matter uncertainty of CE scores obtained

Table 1. Raw dataset

	PG	CT	KL	DG	NK	PŽ	ŠA	BP	MK	BA	RO
I1	9952	32	39	136	270	6	70	613	195	104	124
I2	2370	278	56	148	837	9	7	371	46	285	21
I3	736	13	3	47	46	2	0	36	0	13	5
I4	11327	2220	1658	1379	5303	43	572	2805	923	3955	639
O1	2799	146	110	152	785	16	24	187	54	60	71
O2	13	2	9	2	6	2	2	3	1	2	3

	PL	AN	PV	ŽB	HN	KO	TV	BD	BR	UL	Mean
I1	168	1	353	66	506	224	21	437	1535	429	728
I2	63	9	182	57	467	255	203	532	1324	402	377
I3	6	15	9	10	149	53	227	115	100	9	76
I4	587	316	1601	625	2932	5557	3434	3528	7074	2719	2819
O1	16	17	96	20	450	409	214	449	540	261	327
O2	1	1	1	1	3	3	2	4	4	2	3

Notes: AN – Andrijevica; BA – Berane; BD – Budva; BR – Bar; BP – Bijelo Polje; CT – Cetinje; DG – Danilovgrad; HN – Herceg Novi; KL – Kolašin; KO – Kotor; MK – Mojkovac; NK – Nikšić; PG – Podgorica; PL – Plav; PV – Pljevlja; PŽ – Plužine; RO – Rožaje; ŠA – Šavnik; TV – Tivat; UL – Ulcinj; ŽB – Žabljak.

from inaccurate data. Finally, experts were given the opportunity to choose the extent to which both, self-evaluation $\left(E_f\right)$ and peer-included evaluation score $\left(\Gamma_i\right)$ with resolved uncertainty will be represented in the final DGF result. Depending on the goal of the research, decision-makers are given the possibility to determine the α and β coefficients and thus choose the amount of the self-evaluation and the level of addressing uncertainty. For the purpose of presenting the advantages of the novel methodology in Equation (15), the influence of DEA scores was neglected, and results were obtained relying only on the grey scores $\left(\alpha = 0, \beta = 1\right)$. In that case, final DGF scores were Γ_i scores. The final efficiencies along with corresponding rankings are given in Table 3.

As can be seen in Table 3, according to DEA, due to the small number of DMUs and flexibility in weight assignment, half of the analysed municipalities tended to be efficient, with a score equal to 1 and therefore establishing the final ranking among them is impossible. In addition, the scores of these municipalities and their inherent uncertainty, were further processed with the proposed DGF methodology and all territories were further differentiated. The mean of efficiencies obtained from the DEA model is higher than the both, DGF and CE models, which indicates the neglect of data disturbances. Based on the DGF method, only BR is an efficient territory and remains in the 1st place position of all rankings. The municipality with the lowest efficiency value was determined to be PZ in the case of CE and DGF. However, in the case of DEA, that municipality was KL. To address the volume of aggressive self-evaluation and to calculate the deviation of DEA value scores from CE and DGF values, the Maverick index values were obtained (*MI* in Table 3). The largest Maverick index for DGF and CE was recorded for BD $\left(MI_b = 0.664\right)$ and AN $\left(MI_a = 2.211\right)$, indicating that DEA efficiency value deviates by 66.4% from DGF and 221% from CE. Conclusion can be made that DGF represents a compromise between DEA and CE scores. Scores obtained with the proposed methodology soften the DEA values taking into consideration all of the peers' evaluation efficiencies and

Table 2. Fan–Ma weights

Municipality	PG	CT	KL	DG	NK	PZ	ŠA	BP	MK	BA	RO
Fan–Ma weights	0.032	0.050	0.028	0.059	0.050	0.040	0.033	0.057	0.035	0.059	0.056

Municipality	PL	AN	PV	ŽB	HN	KO	TV	BD	BR	UL	Mean
Fan–Ma weights	0.054	0.038	0.049	0.053	0.056	0.055	0.053	0.058	0.036	0.051	0.048

Table 3. Final scores

	DEA		CE			DGF$^{\#}$				
	Score	*Rank*	*Score*	*Rank*	$MI_a^{\#\#}$	*Score*	*Rank*	$MI_b^{\#\#\#}$	*P*	*Y*
BR	1.000	1	0.805	1	0.242	1.000	1	0.452	98.81	7.26
MK	1.000	1	0.784	3	0.275	0.940	2	0.166	95.24	6.66
BA	1.000	1	0.794	2	0.259	0.939	3	–0.434	90.48	6.31
ZB	1.000	1	0.688	4	0.453	0.866	4	0.225	85.71	6.07
PL	1.000	1	0.687	5	0.456	0.860	5	–0.005	80.95	5.87
TV	1.000	1	0.639	6	0.566	0.814	6	–0.424	76.19	5.71
KO	1.000	1	0.575	8	0.738	0.712	7	0.595	71.43	5.56
HN	1.000	1	0.575	9	0.738	0.699	8	0.128	66.67	5.43
PV	1.000	1	0.576	7	0.738	0.697	9	0.064	61.90	5.30
PG	1.000	1	0.554	10	0.806	0.689	10	0.064	57.14	5.18
SA	1.000	1	0.453	13	1.206	0.627	11	–0.429	52.38	5.06
BP	0.698	15	0.527	11	0.325	0.618	12	0.163	47.62	4.94
BD	0.820	12	0.504	12	0.626	0.613	13	0.664	42.86	4.82
DG	0.683	16	0.432	14	0.580	0.558	14	0.435	38.10	4.70
UL	0.764	14	0.386	15	0.978	0.551	15	0.155	33.33	4.57
CT	0.605	17	0.363	16	0.668	0.519	16	0.430	28.57	4.43
NK	0.493	18	0.317	17	0.554	0.495	17	0.404	23.81	4.29
AN	0.780	13	0.243	18	2.211	0.469	18	0.229	19.05	4.13
RO	0.241	19	0.148	19	0.632	0.423	19	0.337	14.29	3.93
KL	0.229	21	0.093	20	1.456	0.404	20	0.000	9.52	3.69
PZ	0.232	20	0.082	21	1.817	0.402	21	0.386	4.76	3.33

Notes: $^{\#}$DGF: α = 0, β = 1 in Equation (15); $^{\#\#}MI_a = \frac{Ef_{DEA} - Ef_{CE}}{Ef_{CE}}$; $^{\#\#\#}MI_a = \frac{Ef_{DEA} - Ef_{DGF}}{Ef_{DGF}}$.

retained a self-evaluation to some extent. The results are most consistent for the municipalities at the bottom half of the ranking scale, and the rankings of the municipalities described as efficient via DEA are linked with large deviations, which indicate uncertainty in road safety evaluation in the context of DEA.

4.2. Benchmarking

The proposed method of this study has tried to take the advantages of single DEA and GRA methodologies to cover some of the shortcomings and produce more reliable results, which are compatible with practical applications. In real-world applications, the uncertainty is an inevitable feature of results obtained from imprecise and inaccurate data and without considering it, conclusions can be wrong. Efficiency measurement is a process through which useful information can be obtained on how to perform effectively by reinforcing positive behaviours and eliminating inappropriate and unnecessary behaviours in a system. The proposed DGF method aimed to generate more reliable efficiency scores of territories under analysis and in addition, efficient municipalities obtained via DEA become inefficient with DGF and lose their leading and benchmark role. To find benchmarks, it is more realistic to determine a set of comparable territories and target the best performing among them than to make a comparison between territories placed into a single group. Since differences between road safety backgrounds always exist, we propose grouping as described by Chen *et al.* (2016). 1st, the distributions of the DGF score, frequencies, cumulative frequencies, percentiles, and the corresponding probit are determined. A regression model is developed with the DGF as a dependent value and the probit Y as an independent value, with a and b as regression parameters:

$$DGF = a + b \cdot Y. \quad (16)$$

Grouping is performed with a reasonable number of classes, and the class interval is calculated as a critical value of DGF^*, derived with a regression equation according to corresponding probit Y^*, percentiles P^*(shown on the right side of Table 3) and regression equation is determined as:

$$DGF^* = -0.45616 + 0.21405 \cdot Y^*. \quad (14)$$

Finally, municipalities are grouped into 5 classes (Table 4). Within each class, the best performing (territories with the highest DGF value) represent the benchmark municipalities. To ensure that grouping is statistically significant ($P < 0.01$), analysis of variance is performed to verify the obtained results.

5. Discussion

When evaluating the road safety of the territory, the data can be accurate or inaccurate, precise or imprecise, etc., and obtained results can also be either rational and reliable or not. So, addressing uncertainty in such environment without losing any valuable information is crucial and the DGF method represent a modest attempt to do so. In this section, several methods are used to validate the results of the proposed hybrid methodology that integrates DEA with weighted GRA and, to verify suggested benchmarking. The following approaches, which have been used in the literature, have been applied:

- ***ranking-DEA***. In this model, efficiencies are obtained as a self-evaluation. Weights are assigned for each DMU in order to maximize its efficiency score;
- ***ranking-CE***. A CE matrix is used, and the final scores are obtained based on the average of all scores derived from the optimal weights of peer DMUs;
- ***ranking-wDEA***. DEA gives researchers the opportunity to set their own weights and, in this model, constraints related to equal weights for inputs, as well as other equal weights for outputs, were added;
- ***ranking-G/E***. Here, the proposed hybrid methodology is implemented based on data from Si & Ma (2019) and comparisons made with rankings obtained in their research based on the grey correlation degree and relative entropy;
- ***ranking-B/W***. For 31 provinces evaluated with the B/W methodology for evaluating road safety based on DEA in Omrani *et al.* (2020), DGF scores are obtained, and ranking is compared.
- ***ranking-PCA***. The ranking based on PCA of previously mentioned provinces that is given in Omrani *et al.* (2020) is compared to ranking related to the proposed DEA-based weighted GRA methodology applied to the same data (Table B in Appendix B).

To observe correlations between rankings related to 6 presented methods, Pearson's correlation coefficient was obtained, and Spearman's test was performed (Table 5). The highest correlation coefficient was recorded when compared with CE-Ranking, which is not surprise since it

Table 4. Benchmark municipalities

Group	Road safety state	Percentiles P^*	Probit Y^*	$DGF^{*\#}$	Municipality	Benchmark
I	very highly safe territory	96.407	6.8	0.970	BR, MK, BA	BR
II	highly safe territory	72.525	5.6	0.752	ZA, PL, TV	ZA
III	moderate safe territory	27.425	4.4	0.533	KO, HN, PV, PG, SA, BP, BD, DG, UL	KO
IV	low safety territory	3.593	4.2	0.314	CT, NK, AN, RO, KL, PZ	CT
V	very low safety territory	<3.593	<3.2	<0.314	–	–

Note: $^{\#}$at 0.05 significance level, the regression equation is statistically significant $\left(P = 0.000 < 0.001,\ R^2 = 0.968\right)$.

works on similar basis. However, the uncertainty of scores, that is most reflected in higher-ranking municipalities, is addressed with the DGF. This is important since these territories represent the best practices and those with poor performance learn from their strategies in an attempt to improve road safety. With DGF some territories have lost their rank and some improved their position and, a lower deviation index of DGF compared to CE validates the use of DGF instead of CE for purpose to retain peculiarity of DEA. Furthermore, the high degree of harmony related to CE and also B/W-Ranking and raw DEA-Ranking verify the robustness and effectiveness of the proposed methodology and, moreover, indicate the feasibility and convenience of using the proposed methodology for evaluations of the road safety of a territory. Another advantage of the proposed methodology is it can serve to distinguish efficient DMUs, which occurs when number of DMUs and total number of input and output are not in balance (Alpers' rule) and to address uncertainty of scores derived from imprecise data.

The disadvantage of the proposed methodology is that the benchmark entities, territories in our case, lose their leading role so it is necessary to conduct a novel grouping and abstraction of the territory that represents the best practise within each group. The essence of benchmarking is to learn from better performing territories in similar road safety situations, ones that already obtained road safety practices that separates them from the group. In this paper, 21 municipalities are grouped by the state of road safety. 5 classes of municipalities were formed, and the statistical significance of the grouping is verified by analysis of variance (Table 6). The results $(P = 0.000 < 0.05)$ show the acceptability and verification of the proposed grouping and the validation of the derived benchmarks.

4 territories are singled out as leading within each group, and those territories are defined as benchmarks. 3 territories (BR, MK, and BA) are characterised as having very high road safety, 3 (ZA, PL, and TV) as having high road safety, most territories (KO, HN, PV, PG, SA, BP, BD, DG, and UL) as having moderate road safety, 6 territories (CT, NK, AN, RO, KL, and PZ) as having low road safety, and none as having extremely low road safety. On the other hand, it is important to emphasize that, according to all 3 methods, DEA, CE, and DGF, the same 3 territories were identified as having the lowest level of traffic safety, indicating that decision-makers have clearly defined territories that require an urgent implementation of measures and investments in order to improve traffic safety. When it comes to the best territories in terms of the level of traffic safety, the DGF method more clearly classifies the territories, with the 4 best territories identified by both the CE method and the DGF method. This is also important for decision-makers with regard to the potential rewarding and highlighting of the best territories as an example of best practices in road safety.

6. Conclusions

It is very important to adequately address issues of road accidents and traffic injuries, as they remain great problems for the public health system. The aggregation of different variables into one road safety performance index is a popular concept in evaluating road safety and in comparing the performance of entities, and DEA seems to be useful technique for that purpose. However, in the conventional DEA models, the uncertainty of data, which is an undeniable feature of real-world problems is ignored; hence the results will be unreliable and considering their uncertainty is essential. Likewise, safety indicators used for index construction are not always precise and accurate so the composite index obtained with such data is considered uncertain and can be misleading. Hence, to reflect the inaccuracy of data, address uncertainty of results obtained from such kind of data, provide robust solutions, and to cover some of the shortcomings of the DEA, the score based on the DGF method for road safety composite index construction was developed in this study. Since GRA has proven superiority to handle uncertainty for evaluation of a system when the decision-makers are not certain about the sufficiency and completeness of data, the authors of this paper have proposed its integration with DEA in order to give a more reliable picture of road safety.

Table 5. Ranking correlations between different methods

	Ranking-DEA	Ranking-CE	Ranking-wDEA	Ranking-G/E	Ranking-B/W	Ranking- PCA
Ranking-DGF						
Pearson's correlation	0.812#	0.991#	0.632#	0.559#	0.847#	0.689#
Significance (2-tailed)	0.000	0.000	0.002	0.059	0.000	0.000
N	21	21	21	12	31	31

Note: #correlation is significant at the 0.01 level (2-tailed).

Table 6. Analysis of variance

	Sum of squares	*df*	Mean square	*F*	Significance (*P*)
Between groups	0.637	4	0.159	54.650	0.000
Within groups	0.047	16	0.003	–	–
Total	0.684	20	–	–	–

The main contribution of this paper is the proposed extension of the DEA method based on GRA that are used to evaluate the efficiency of DMUs having imprecise inputs and outputs. This novel methodology has practical and simple usefulness to assess road safety situation of the territory, to obtain reliable safety composite index and overcome its initial uncertainty. The proposed approach in addition to providing acceptable results avoids time-consuming computations and consequently reduces the solution time. Furthermore, the proposed methodology can also serve to further distinguish efficient DMUs with reducing self-evaluation including peers in assessment but also retaining the peculiarity of DEA at the same time. The DGF model can be applied to identify inefficient territories in a more reasonable and rationale way, as it provides a wider range of efficiency assessment when inaccurate data are involved. Policymakers and road safety experts can evaluate the territories' efficiency and make important decisions to improve the current safety situation.

Step of the DGF methodology are as follows: 1st, based on the literature and availability, indicators were selected, normalized, and categorised as inputs and outputs. A numerical example in this paper is applied to a case study of Montenegro with behaviour-related indicators as inputs and final road safety outcomes as outputs. Afterwards, DEA was used to calculate municipalities' efficiencies and to construct a CE matrix. Fan–Ma weights were derived in order to maintain the essence of DEA and the grey relational coefficient that represents the similarity between efficiencies of DMUs with peers' weights and an ideal reference set, was obtained next. Finally, the grey relational degree was calculated with the final DGF score derivation. The final scores represent softened DEA values, taking into account all the peers' evaluation efficiencies, and simultaneously maintain self-evaluation to some extent. The DGF method provides an enriched picture compared to the basic DEA approach, and both, efficient and inefficient municipalities are further differentiated. Moreover, based on the DGF results, the territories were further ranked and classified into 5 classes according to the road safety level from best to worst. In addition, for each road safety group, the best-performing municipality is identified and determined as a benchmark.

One of the advantages of the proposed methodology is that it does not require complicated software and also can be used when the number of road safety indicators is higher than the number of territories under evaluation because the final results are not affected by the number of DMUs. This novel method can be treated as a unique methodology for annual road safety monitoring and measuring achievements of implemented measures. The proposed method can be applied to many circumstances with inaccurate statistics so that they can be adapted for a wider range in efficiency evaluation. In addition, it can be considered as a natural extension of DEA and CE when input data are considered inaccurate and imprecise, and the verified robustness of the results indicates the feasibility of addressing uncertainty of obtained DEA/CE scores, in many other fields besides road safety. The DGF method is easily applied to a small sample and when total number of indictors is not in respect with number of territories under evaluation. Furthermore, with the proposed approach, full and rational ranking is derived under uncertain environment and obtained score is verified indicating that DGF is superior to basic DEA as regards the territory ranking.

However, given that GRA is normalization-based methodology, the results might vary based on the type of normalization, which is the largest shortcoming of this method. On the other hand, the proposed GRA has encountered the sensitivity problem arising from the parameter setup of the distinguish coefficient. The different distinguish coefficients may lead to different solution results so in the future work, researchers should try several different distinguishing coefficients and analyse the impact on the results. One of limitation of this study is that DEA might have multiple weight solutions resulting in different values of peer efficiencies. Future researchers could investigate the issue of alternative optimal solutions in their proposed model. One more disadvantage is the fact that when uncertainty of DEA scores is addressed some of DMUs lose their rank and cannot represent the best practice so, the novel benchmarking process is needed and in future work other grouping and clustering methods should be investigated. Since the obtained efficiencies depend on chosen input and output indicators, in future research, sensitivity analysis should be implemented in order to define standardized categories of indicators that can best describe the level of road safety of territories. To increase the success of the proposed methodology, territory evaluation can be expanded considering an interval or data placed in a hierarchical multi-layered structure. Also, different weighting methods for calculating grey relational degree, when larger sample is the case, should be considered.

Appendix A

Table A1. CE matrix derived by DEA

DMU	Target DMU																				
	PG	CT	KL	DG	NK	PZ	ŠA	BP	MK	BA	RO	PL	AN	PV	ŽB	HN	KO	TV	BD	BR	UL
PG	1.000	0.497	0.061	0.982	0.497	0.427	0.227	0.577	0.449	0.616	0.580	0.557	0.211	0.545	0.359	1.000	0.511	0.507	1.000	0.525	0.499
CT	0.056	0.605	0.231	0.383	0.605	0.228	0.205	0.342	0.161	0.425	0.331	0.321	0.266	0.174	0.408	0.355	0.591	0.568	0.388	0.469	0.511
KL	0.006	0.078	0.229	0.075	0.078	0.069	0.101	0.130	0.023	0.119	0.128	0.126	0.099	0.122	0.132	0.072	0.088	0.073	0.076	0.040	0.092
DG	0.222	0.370	0.138	0.683	0.370	0.498	0.375	0.512	0.336	0.570	0.509	0.499	0.394	0.361	0.460	0.676	0.405	0.397	0.670	0.251	0.382
NK	0.106	0.493	0.102	0.339	0.493	0.147	0.148	0.367	0.154	0.383	0.362	0.344	0.155	0.283	0.278	0.308	0.484	0.473	0.337	0.429	0.475
PZ	0.010	0.012	0.041	0.171	0.012	0.232	0.127	0.081	0.042	0.101	0.081	0.083	0.130	0.053	0.102	0.190	0.016	0.012	0.184	0.034	0.012
ŠA	0.183	0.079	0.362	0.265	0.079	0.342	1.000	0.818	0.712	0.521	0.814	0.842	1.000	0.482	1.000	0.260	0.180	0.169	0.230	0.023	0.160
BP	0.232	0.513	0.228	0.663	0.513	0.444	0.460	0.698	0.266	0.697	0.697	0.678	0.433	0.682	0.553	0.633	0.534	0.516	0.663	0.427	0.529
MK	1.000	0.434	0.259	1.000	0.434	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.711	0.946	1.000	0.552	0.545	0.901	0.171	0.522
BA	0.102	1.000	1.000	0.714	1.000	0.556	0.710	1.000	0.247	1.000	0.982	0.957	0.731	0.764	1.000	0.671	1.000	0.917	0.718	0.611	1.000
RO	0.035	0.089	0.137	0.137	0.089	0.127	0.197	0.242	0.064	0.205	0.241	0.238	0.183	0.253	0.219	0.132	0.121	0.113	0.131	0.033	0.120
PL	0.127	0.315	0.557	0.901	0.315	0.869	1.000	1.000	0.279	1.000	1.000	1.000	0.926	1.000	0.974	0.890	0.363	0.319	0.913	0.339	0.333
AN	0.005	0.127	0.282	0.330	0.127	0.498	0.147	0.102	0.358	0.210	0.095	0.098	0.780	0.028	0.723	0.343	0.198	0.183	0.330	0.049	0.085
PV	0.199	0.864	0.253	0.480	0.864	0.239	0.308	0.743	0.208	0.657	0.742	0.701	0.283	1.000	0.481	0.434	0.841	0.815	0.474	0.594	0.906
ŽB	0.145	0.328	0.474	1.000	0.328	1.000	1.000	0.926	0.356	1.000	0.920	0.926	1.000	0.684	1.000	1.000	0.386	0.349	1.000	0.280	0.347
HN	0.511	0.546	0.099	1.000	0.546	0.532	0.319	0.637	0.486	0.701	0.636	0.614	0.320	0.508	0.463	1.000	0.572	0.567	1.000	0.473	0.550
KO	0.209	0.853	0.206	0.539	0.853	0.246	0.293	0.737	0.344	0.723	0.725	0.691	0.318	0.511	0.589	0.486	1.000	1.000	0.505	0.277	0.978
TV	0.055	0.837	0.243	1.000	0.837	0.814	0.220	0.331	1.000	0.601	0.313	0.307	0.786	0.112	1.000	0.999	1.000	1.000	1.000	0.337	0.622
BD	0.317	0.490	0.119	0.819	0.490	0.479	0.333	0.579	0.380	0.640	0.577	0.560	0.341	0.444	0.457	0.806	0.515	0.507	0.820	0.417	0.493
BR	0.473	1.000	0.199	1.000	1.000	0.446	0.424	1.000	0.441	0.999	1.000	0.954	0.401	1.000	0.666	0.929	1.000	0.982	1.000	1.000	1.000
UL	0.102	0.754	0.158	0.234	0.754	0.093	0.128	0.426	0.103	0.357	0.425	0.394	0.117	0.712	0.245	0.206	0.657	0.632	0.231	0.622	0.764
Reference set of normalized matrix	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A2. Grey correlation coefficient

DMU	Target DMU																				
	PG	CT	KL	DG	NK	PZ	ŠA	BP	MK	BA	RO	PL	AN	PV	ŽB	HN	KO	TV	BD	BR	UL
PG	1.000	0.495	0.338	0.962	0.495	0.449	0.368	0.521	0.470	0.539	0.523	0.508	0.364	0.517	0.412	1.000	0.501	0.501	1.000	0.507	0.496
CT	0.345	0.556	0.384	0.429	0.556	0.376	0.361	0.411	0.368	0.439	0.407	0.403	0.380	0.370	0.432	0.419	0.546	0.534	0.430	0.479	0.502
KL	0.334	0.349	0.383	0.333	0.349	0.333	0.333	0.346	0.333	0.338	0.345	0.344	0.333	0.356	0.341	0.333	0.350	0.347	0.333	0.337	0.352
DG	0.390	0.439	0.357	0.593	0.439	0.481	0.418	0.485	0.424	0.511	0.483	0.478	0.426	0.432	0.454	0.589	0.453	0.450	0.584	0.395	0.444
NK	0.357	0.493	0.348	0.412	0.493	0.353	0.345	0.420	0.366	0.421	0.419	0.411	0.348	0.404	0.383	0.402	0.488	0.484	0.411	0.461	0.485
PZ	0.334	0.333	0.333	0.358	0.333	0.377	0.340	0.333	0.338	0.333	0.333	0.333	0.341	0.339	0.333	0.364	0.333	0.333	0.361	0.336	0.333
ŠA	0.378	0.349	0.429	0.386	0.349	0.415	1.000	0.716	0.629	0.484	0.711	0.743	1.000	0.484	1.000	0.385	0.375	0.373	0.375	0.333	0.370
BP	0.393	0.503	0.383	0.579	0.503	0.456	0.454	0.603	0.400	0.598	0.603	0.587	0.443	0.605	0.501	0.558	0.514	0.505	0.579	0.460	0.512
MK	1.000	0.466	0.393	1.000	0.466	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.627	0.893	1.000	0.524	0.520	0.823	0.371	0.508
BA	0.356	1.000	1.000	0.618	1.000	0.512	0.608	1.000	0.394	1.000	0.962	0.914	0.626	0.673	1.000	0.585	1.000	0.856	0.621	0.556	1.000
RO	0.340	0.352	0.357	0.349	0.352	0.348	0.359	0.377	0.343	0.361	0.377	0.376	0.355	0.394	0.365	0.348	0.359	0.358	0.347	0.336	0.360
PL	0.363	0.419	0.520	0.824	0.419	0.780	1.000	1.000	0.404	1.000	1.000	1.000	0.858	1.000	0.945	0.808	0.436	0.420	0.842	0.425	0.426
AN	0.333	0.361	0.400	0.408	0.361	0.481	0.345	0.338	0.432	0.363	0.337	0.337	0.671	0.333	0.618	0.414	0.380	0.377	0.408	0.339	0.351
PV	0.383	0.784	0.391	0.471	0.784	0.379	0.394	0.641	0.381	0.567	0.641	0.605	0.386	1.000	0.464	0.451	0.756	0.728	0.468	0.546	0.839
ŽB	0.368	0.424	0.477	1.000	0.424	1.000	1.000	0.862	0.431	1.000	0.852	0.860	1.000	0.606	1.000	1.000	0.445	0.431	1.000	0.404	0.431
HN	0.504	0.521	0.347	1.000	0.521	0.499	0.398	0.559	0.487	0.601	0.558	0.543	0.399	0.497	0.456	1.000	0.535	0.533	1.000	0.481	0.523
KO	0.386	0.771	0.377	0.501	0.771	0.382	0.389	0.636	0.427	0.619	0.626	0.597	0.398	0.498	0.522	0.475	1.000	0.999	0.483	0.403	0.957
TV	0.345	0.752	0.388	1.000	0.752	0.715	0.366	0.407	1.000	0.530	0.401	0.398	0.678	0.354	1.000	0.997	1.000	1.000	1.000	0.424	0.566
BD	0.422	0.492	0.353	0.719	0.492	0.472	0.403	0.522	0.441	0.555	0.521	0.510	0.406	0.466	0.453	0.705	0.503	0.500	0.719	0.456	0.493
BR	0.485	1.000	0.374	1.000	1.000	0.457	0.438	1.000	0.466	0.999	1.000	0.908	0.429	1.000	0.573	0.868	1.000	0.964	1.000	1.000	1.000
UL	0.356	0.667	0.363	0.376	0.667	0.339	0.340	0.445	0.353	0.411	0.444	0.431	0.338	0.627	0.373	0.369	0.589	0.573	0.375	0.564	0.677

Table A3. Weighted grey matrix

DMU	Target DMU																				
	PG	CT	KL	DG	NK	PZ	ŠA	BP	MK	BA	RO	PL	AN	PV	ŽB	HN	KO	TV	BD	BR	UL
PG	0.032	0.025	0.009	0.056	0.025	0.018	0.012	0.030	0.016	0.032	0.029	0.028	0.014	0.025	0.022	0.056	0.027	0.026	0.058	0.018	0.025
CT	0.011	0.028	0.011	0.025	0.028	0.015	0.012	0.023	0.013	0.026	0.023	0.022	0.015	0.018	0.023	0.024	0.030	0.028	0.025	0.017	0.025
KL	0.011	0.017	0.011	0.020	0.017	0.013	0.011	0.020	0.012	0.020	0.019	0.019	0.013	0.017	0.018	0.019	0.019	0.018	0.019	0.012	0.018
DG	0.013	0.022	0.010	0.035	0.022	0.019	0.014	0.028	0.015	0.030	0.027	0.026	0.016	0.021	0.024	0.033	0.025	0.024	0.034	0.014	0.023
NK	0.012	0.024	0.010	0.024	0.024	0.014	0.011	0.024	0.013	0.025	0.024	0.022	0.013	0.020	0.020	0.023	0.027	0.025	0.024	0.017	0.025
PZ	0.011	0.017	0.009	0.021	0.017	0.015	0.011	0.019	0.012	0.020	0.019	0.018	0.013	0.017	0.018	0.021	0.018	0.018	0.021	0.012	0.017
ŠA	0.012	0.017	0.012	0.023	0.017	0.017	0.033	0.041	0.022	0.028	0.040	0.040	0.038	0.024	0.053	0.022	0.020	0.020	0.022	0.012	0.019
BP	0.013	0.025	0.011	0.034	0.025	0.018	0.015	0.034	0.014	0.035	0.034	0.032	0.017	0.030	0.027	0.031	0.028	0.027	0.033	0.016	0.026
MK	0.032	0.023	0.011	0.059	0.023	0.040	0.033	0.057	0.035	0.059	0.056	0.054	0.038	0.031	0.047	0.056	0.029	0.027	0.047	0.013	0.026
BA	0.012	0.050	0.028	0.036	0.050	0.021	0.020	0.057	0.014	0.059	0.054	0.049	0.024	0.033	0.053	0.033	0.055	0.045	0.036	0.020	0.051
RO	0.011	0.017	0.010	0.020	0.017	0.014	0.012	0.022	0.012	0.021	0.021	0.020	0.014	0.019	0.019	0.020	0.020	0.019	0.020	0.012	0.018
PL	0.012	0.021	0.014	0.048	0.021	0.031	0.033	0.057	0.014	0.059	0.056	0.054	0.033	0.049	0.050	0.046	0.024	0.022	0.048	0.015	0.022
AN	0.011	0.018	0.011	0.024	0.018	0.019	0.011	0.019	0.015	0.021	0.019	0.018	0.026	0.016	0.033	0.023	0.021	0.020	0.023	0.012	0.018
PV	0.012	0.039	0.011	0.028	0.039	0.015	0.013	0.037	0.013	0.033	0.036	0.033	0.015	0.049	0.025	0.025	0.041	0.038	0.027	0.020	0.043
ŽB	0.012	0.021	0.013	0.059	0.021	0.040	0.033	0.049	0.015	0.059	0.048	0.047	0.038	0.030	0.053	0.056	0.024	0.023	0.058	0.014	0.022
HN	0.016	0.026	0.010	0.059	0.026	0.020	0.013	0.032	0.017	0.035	0.031	0.029	0.015	0.024	0.024	0.056	0.029	0.028	0.058	0.017	0.027
KO	0.013	0.038	0.010	0.029	0.038	0.015	0.013	0.036	0.015	0.036	0.035	0.032	0.015	0.024	0.028	0.027	0.055	0.053	0.028	0.014	0.048
TV	0.011	0.037	0.011	0.059	0.037	0.029	0.012	0.023	0.035	0.031	0.023	0.022	0.026	0.017	0.053	0.056	0.055	0.053	0.058	0.015	0.029
BD	0.014	0.024	0.010	0.042	0.024	0.019	0.013	0.030	0.015	0.033	0.029	0.028	0.016	0.023	0.024	0.040	0.028	0.026	0.041	0.016	0.025
BR	0.016	0.050	0.010	0.059	0.050	0.018	0.014	0.057	0.016	0.059	0.056	0.049	0.016	0.049	0.030	0.049	0.055	0.051	0.058	0.036	0.051
UL	0.012	0.033	0.010	0.022	0.033	0.014	0.011	0.025	0.012	0.024	0.025	0.023	0.013	0.031	0.020	0.021	0.032	0.030	0.022	0.020	0.034

Appendix B

Table B. Ranking

DMU	Ranking-DGF	Ranking-B/W	Ranking-PCA	Ranking-G/E
DMU1	2	21	27	1
DMU2	20	19	24	2
DMU3	13	18	17	3
DMU4	9	8	2	4
DMU5	11	13	23	5
DMU6	10	12	18	6
DMU7	19	10	8	7
DMU8	1	1	1	8
DMU9	12	15	13	9
DMU10	6	9	7	10
DMU11	7	5	6	11
DMU12	26	30	26	12
DMU13	15	14	9	–
DMU14	25	25	21	–
DMU15	4	7	11	–
DMU16	3	3	10	–
DMU17	23	20	14	–
DMU18	17	16	16	–
DMU19	14	11	12	–
DMU20	24	23	28	–
DMU21	8	4	5	–
DMU22	29	29	29	–
DMU23	31	31	31	–
DMU24	21	26	19	–
DMU25	30	28	30	–
DMU26	27	27	22	–
DMU27	22	22	20	–
DMU28	28	24	25	–
DMU29	5	2	3	–
DMU30	16	17	15	–
DMU31	18	6	4	–

Author contributions

Boris Antić conceived the study, *Mirjana Grdinić-Rakonjac* and *Vladimir Pajković* were responsible for the design and development of the data analysis.

Mirjana Grdinić-Rakonjac developed the methodology, *Mirjana Grdinić-Rakonjac* and *Vladimir Pajković* were responsible for data collection and analysis.

Vladimir Pajković were responsible for data interpretation.

Mirjana Grdinić-Rakonjac wrote the first draft of the article, *Boris Antić* and *Vladimir Pajković* were responsible for editing.

All authors have read and agreed to the published version of the manuscript.

Disclosure statement

Authors do not have any competing financial, professional, or personal interests from other parties.

References

Al-Haji, G. 2007. *Road Safety Development Index (RSDI): Theory, Philosophy and Practice*. Dissertation No 1100. Linköping University, Norrköping, Sweden. 159 p. Available from Internet: https://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-8812

Alper, D.; Sinuany-Stern, Z.; Shinar, D. 2015. Evaluating the efficiency of local municipalities in providing traffic safety using the data envelopment analysis, *Accident Analysis & Prevention* 78: 39–50. https://doi.org/10.1016/j.aap.2015.02.014

Amini, M.; Dabbagh, R.; Omrani, H. 2019. A fuzzy data envelopment analysis based on credibility theory for estimating road safety, *Decision Science Letters* 8: 275–284. https://doi.org/10.5267/j.dsl.2019.1.001

Anderson, T. R.; Hollingsworth, K.; Inman, L. 2002. The fixed weighting nature of a cross-evaluation model, *Journal of Productivity Analysis* 17(3): 249–255. https://doi.org/10.1023/A:1015012121760

Antić, B.; Grdinić, M.; Pešić, D.; Pajković, V. 2020. Benchmarking of the road safety performance among the regions by using DEA, *Transportation Research Procedia* 45: 78–86. https://doi.org/10.1016/j.trpro.2020.02.065

Ayati, E.; Pirayesh Neghab, M. A.; Sadeghi, A. A.; Mohammadzadeh Moghaddam, A. 2012. Introducing roadside hazard severity indicator based on evidential reasoning approach, *Safety Science* 50(7): 1618–1626. https://doi.org/10.1016/j.ssci.2012.03.023

Babaee, S.; Bagherikahvarin, M.; Sarrazin, R.; Shen, Y.; Hermans, E. 2015. Use of DEA and PROMETHEE II to assess the performance of older drivers, *Transportation Research Procedia* 10: 798–808. https://doi.org/10.1016/j.trpro.2015.09.033

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: 106993. https://doi.org/10.1016/j.cie.2020.106993

Baležentis, A.; Baležentis, T. 2011. Assessing the efficiency of Lithuanian transport sector by applying the methods of MULTIMOORA and data envelopment analysis, *Transport* 26(3): 263–270. https://doi.org/10.3846/16484142.2011.621146

Banker, R. D.; Charnes, A.; Cooper, W. W. 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science* 30(9): 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

Bao, Q.; Ruan, D.; Shen, Y.; Hermans, E.; Janssens, D. 2012. Improved hierarchical fuzzy TOPSIS for road safety performance evaluation, *Knowledge-Based Systems* 32: 84–90. https://doi.org/10.1016/j.knosys.2011.08.014

Bastos, J. T.; Shen, Y.; Hermans, E.; Brijs, T.; Wets, G.; Ferraz, A. C. P. 2015. Traffic fatality indicators in Brazil: state diagnosis based on data envelopment analysis research, *Accident Analysis & Prevention* 81: 61–73. https://doi.org/10.1016/j.aap.2015.01.024

Behdani, Z.; Darehmiraki, M. 2022. An alternative approach to rank efficient DMUs in DEA via cross-efficiency evaluation, Gini coefficient, and Bonferroni mean, *Journal of the Operations Research Society of China* 10(4): 763–783. https://doi.org/10.1007/s40305-019-00264-x

Behnood, H. R.; Ayati, E.; Brijs, T.; Pirayesh Neghab, M.; Shen, Y. 2017. A fuzzy decision-support system in road safety planning, *Proceedings of the Institution of Civil Engineers – Transport* 170(5): 305–317. https://doi.org/10.1680/jtran.15.00062

Blagojević, A.; Vesković, S.; Kasalica, S.; Gojić, A.; Allamani, A. 2020. The application of the fuzzy AHP and DEA for measuring the efficiency of freight transport railway undertakings, *Operational Research in Engineering Sciences: Theory and Applications* 3(2): 1–23. Available from Internet: https://oresta.org/menu-script/index.php/oresta/article/view/52

Castro-Nuño, M.; Arévalo-Quijada, M. T. 2018. Assessing urban road safety through multidimensional indexes: Application of multicriteria decision making analysis to rank the Spanish provinces, *Transport Policy* 68: 118–129. https://doi.org/10.1016/j.tranpol.2018.04.017

Charnes, A.; Cooper, W. W.; Rhodes, E. 1978. Measuring the efficiency of decision making units, *European Journal of Operational Research* 2(6): 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Chen, F.; Wang, J.; Deng, Y. 2015. Road safety risk evaluation by means of improved entropy TOPSIS–RSR, *Safety Science* 79: 39–54. https://doi.org/10.1016/j.ssci.2015.05.006

Chen, F.; Wu, J.; Chen, X.; Wang, J.; Wang, D. 2016. Benchmarking road safety performance: identifying a meaningful reference (best-in-class), *Accident Analysis & Prevention* 86: 76–89. https://doi.org/10.1016/j.aap.2015.10.018

Deng, J. 1989. Introduction to Grey system theory, *The Journal of Grey System* 1(1): 1–24.

Dutta, P.; Jaikumar, B.; Arora, M. S. 2022. Applications of data envelopment analysis in supplier selection between 2000 and 2020: a literature review, *Annals of Operations Research* 315(2): 1399–1454. https://doi.org/10.1007/s10479-021-03931-6

Đalić, I.; Stević, Ž.; Ateljević, J.; Turskis, Z.; Zavadskas, E. K.; Mardani, A. 2021. A novel integrated MCDM-SWOT-TOWS model for the strategic decision analysis in transportation company, *Facta Universitatis: Mechanical Engineering* 19(3): 401–422. https://doi.org/10.22190/fume201125032d

Ehrgott, M.; Holder, A.; Nohadani, O. 2018. Uncertain data envelopment analysis, *European Journal of Operational Research* 268(1): 231–242. https://doi.org/10.1016/j.ejor.2018.01.005

ETSC. 2001. *Transport Safety Performance Indicators*. European Transport Safety Council (ETSC), Brussels, Belgium. 56 p. Available from Internet: https://etsc.eu/transport-safety-performance-indicators

Fan, Z.-P.; Zheng, L.-H.; Pan, D.-H. 1996. A Subjective and objective synthetic method for multiple attribute decision making, *IFAC Proceedings Volumes* 29(1): 5635–5640.
https://doi.org/10.1016/S1474-6670(17)58580-8

Fancello, G.; Carta, M.; Serra, P. 2020. Data envelopment analysis for the assessment of road safety in urban road networks: a comparative study using CCR and BCC models, *Case Studies on Transport Policy* 8(3): 736–744.
https://doi.org/10.1016/j.cstp.2020.07.007

Folla, K.; Nikolaou, P.; Dimitriou, L.; Yannis, G. 2021. Benchmarking analysis of road safety levels for an extensive and representative dataset of European cities, *Advances in Intelligent Systems and Computing* 1278: 1066–1075.
https://doi.org/10.1007/978-3-030-61075-3_102

Geerlings, H.; Klementschitz, R.; Mulley, C. 2006. Development of a methodology for benchmarking public transportation organisations: a practical tool based on an industry sound methodology, *Journal of Cleaner Production* 14(2): 113–123.
https://doi.org/10.1016/j.jclepro.2005.03.021

Gitelman, V.; Doveh, E.; Hakkert, S. 2010. Designing a composite indicator for road safety, *Safety Science* 48(9): 1212–1224.
https://doi.org/10.1016/j.ssci.2010.01.011

Grdinić-Rakonjac, M.; Antić, B.; Pešić, D.; Pajković, V. 2021. Construction of road safety composite indicator using grey relational analysis, *Promet – Traffic & Transportation* 33(1): 103–116. Available from Internet: https://traffic2.fpz.hr/index.php/promtt/article/view/2914

Hermans, E. 2009. *A Methodology for Developing a Composite Road Safety Performance Index for Cross-Country Comparison*. PhD Dissertation. Hasselt University, Belgium. 261 p.

Hermans, E.; Van den Bosshe, F.; Wets, G. 2008. Combining road safety information in a performance index, *Accident Analysis & Prevention* 40(4): 1337–1344.
https://doi.org/10.1016/j.aap.2008.02.004

Hu, Q.-Z.; Lu, H.-P.; Deng, W. 2010. Evaluating the urban public transit network based on the attribute recognition model, *Transport* 25(3): 300–306.
https://doi.org/10.3846/transport.2010.37

Jameel, A. K.; Evdorides, H. 2021. Developing a safer road user behaviour index, *IATSS Research* 45(1): 70–78.
https://doi.org/10.1016/j.iatssr.2020.06.006

Javed, S. A.; Khan, A. M.; Dong, W.; Raza, A.; Liu, S. 2019. Systems evaluation through new grey relational analysis approach: an application on thermal conductivity – petrophysical parameters' relationships, *Processes* 7(6): 348.
https://doi.org/10.3390/pr7060348

Jiang, B.; Lio, W.; Li, X. 2019. An uncertain DEA model for scale efficiency evaluation, *IEEE Transactions on Fuzzy Systems* 27(8): 1616–1624. https://doi.org/10.1109/tfuzz.2018.2883546

Kang, L.; Wu, C. 2022. Evaluating the performance of Chinese provincial road safety based on the output–input ratio, *Transportation Letters: the International Journal of Transportation Research* 14(2): 114–123.
https://doi.org/10.1080/19427867.2020.1819077

Koornstra, M.; Lynam, D.; Nilsson, G.; Noordzij, P.; Pettersson, H.-E.; Wegman, F.; Wouters, P. 2002. *SUNflower: a Comparative Study of the Development of Road Safety in Sweden, the United Kingdom, and the Netherlands*. SWOV Institute for Road Safety Research, The Netherlands. 159 p.

Küçükönder, H.; Demirarslan, P. Ç.; Burgut, A.; Boğa, M. 2019. A hybrid approach of data envelopment analysis based grey relational analysis: a study on egg yield, *Pakistan Journal of Zoology* 51(3): 903–912.
https://doi.org/10.17582/journal.pjz/2019.51.3.903.912

Lee, Y.-C. 2019. Ranking DMUs by combining cross-efficiency scores based on Shannon's entropy, *Entropy* 21(5): 467.
https://doi.org/10.3390/e21050467

Liu, B. 2012. Why is there a need for uncertainty theory?, *Journal of Uncertain Systems* 6(1): 3–10. Available from Internet: https://jus.worldscientific.com/pdf/jusVol06No1paper01.pdf

Liu, Y.; Huang, X.; Duan, J.; Zhang, H. 2017. The assessment of traffic accident risk based on grey relational analysis and fuzzy comprehensive evaluation method, *Natural Hazards* 88: 1409–1422. https://doi.org/10.1007/s11069-017-2923-2

Lu, M.; Wevers, K. 2007. Application of grey relational analysis for evaluating road traffic safety measures: advanced driver assistance systems against infrastructure redesign, *IET Intelligent Transport Systems* 1(1): 3–14.
https://doi.org/10.1049/iet-its:20060022

Ma, J.; Fan, Z.-P.; Huang, L.-H. 1999. A subjective and objective integrated approach to determine attribute weights, *European Journal of Operational Research* 112(2): 397–404.
https://doi.org/10.1016/S0377-2217(98)00141-6

Ma, Z.; Shao, C.; Ma, S.; Ye, Z. 2011. Constructing road safety performance indicators using fuzzy Delphi method and grey Delphi method, *Expert Systems with Applications* 38(3): 1509–1514.
https://doi.org/10.1016/j.eswa.2010.07.062

Mahmoudi, R.; Emrouznejad, A.; Shetab-Boushehri, S.-N.; Hejazi S. R. 2020. The origins, development and future directions of data envelopment analysis approach in transportation systems, *Socio-Economic Planning Sciences* 69: 100672.
https://doi.org/10.1016/j.seps.2018.11.009

Mardani, A.; Zavadskas, E. K.; Khalifah, Z.; Jusoh, A.; Nor, K. M. D. 2016. Multiple criteria decision-making techniques in transportation systems: a systematic review of the state of the art literature, *Transport* 31(3): 359–385.
https://doi.org/10.3846/16484142.2015.1121517

Mardani, A.; Zavadskas, E. K.; Štreimikienė, D.; Jusoh, A.; Khoshnoudi, M. 2017. A comprehensive review of data envelopment analysis (DEA) approach in energy efficiency, *Renewable and Sustainable Energy Reviews* 70: 1298–1322.
https://doi.org/10.1016/j.rser.2016.12.030

Markabi, M.; S.; Sabbagh, M. 2014. A hybrid method of grey relational analysis and data envelopment analysis for evaluating and selecting efficient suppliers plus a novel ranking method for grey numbers, *Journal of Industrial Engineering and Management* 7(5): 25. https://doi.org/10.3926/jiem.872

Memiş, S.; Demir, E.; Karamaşa, Ç.; Korucuk, S. 2020. Prioritization of road transportation risks: an application in Giresun province, *Operational Research in Engineering Sciences: Theory and Applications* 3(2): 111–126. Available from Internet: https://oresta.org/menu-script/index.php/oresta/article/view/63

Mirzapour Al-E-Hashem, S. M. J.; Malekly, H.; Aryanezhad, M. B. 2011. A multi-objective robust optimization model for multi-product multi-site aggregate production planning in a supply chain under uncertainty, *International Journal of Production Economics* 134(1): 28–42.
https://doi.org/10.1016/j.ijpe.2011.01.027

Mitrović Simić, J.; Stević, Ž.; Zavadskas, E. K.; Bogdanović, V.; Subotić, M.; Mardani, A. 2020. A novel CRITIC-fuzzy FUCOM-DEA-fuzzy MARCOS model for safety evaluation of road sections based on geometric parameters of road, *Symmetry* 12(12): 2006. https://doi.org/10.3390/sym12122006

Nenadić, D. 2019. Ranking dangerous sections of the road using MCDM model, *Decision Making: Applications in Management and Engineering* 2(1): 115–131.
https://doi.org/10.31181/dmame1901115n

Nikolaou, P.; Dimitriou, L. 2018. Evaluation of road safety policies performance across Europe: results from benchmark analysis for a decade, *Transportation Research Part A*: *Policy and Practice* 116: 232–246. https://doi.org/10.1016/j.tra.2018.06.026

OECD. 2008. *Handbook on Constructing Composite Indicators*: *Methodology and User Guide*. Organisation for Economic Co-Operation and Development (OECD). 162 p. https://doi.org/10.1787/9789264043466-en

Omrani, H.; Amini, M.; Alizadeh, A. 2020. An integrated group best–worst method – data envelopment analysis approach for evaluating road safety: a case of Iran, *Measurement* 152: 107330. https://doi.org/10.1016/j.measurement.2019.107330

Pakkar, M. S. 2016. A hierarchical aggregation approach for indicators based on data envelopment analysis and analytic hierarchy process, *Systems* 4(1): 6. https://doi.org/10.3390/systems4010006

Peykani, P.; Mohammadi, E.; Emrouznejad, A.; Pishvaee, M. S.; Rostamy-Malkhalifeh, M. 2019. Fuzzy data envelopment analysis: an adjustable approach, *Expert Systems with Applications* 136: 439–452. https://doi.org/10.1016/j.eswa.2019.06.039

Qazvini, A. W.; Javanshir, H.; Seyedhosseni, S. M. 2016. An application of fuzzy TOPSIS for identifying and analyzing black spots suburban areas, *Management Science Letters* 6: 561–568. https://doi.org/10.5267/j.msl.2016.6.005

Raheel Shah, S. A.; Ahmad, N.; Shen, Y.; Kamal, M. A.; Basheer, M. A.; Brijs, T. 2019. Relationship between road traffic features and accidents: An application of two-stage decision-making approach for transportation engineers, *Journal of Safety Research* 69: 201–215. https://doi.org/10.1016/j.jsr.2019.01.001

Rassafi, A. A.; Ganji, S. S.; Pourkhani, H. 2018. Road safety assessment under uncertainty using a multi attribute decision analysis based on Dempster–Shafer theory, *KSCE Journal of Civil Engineering* 22(8): 3137–3152. https://doi.org/10.1007/s12205-017-1854-5

René, S. H.; Becker, U.; Manz, H. 2016. Grey systems theory time series prediction applied to road traffic safety in Germany, *IFAC – Papers on Line* 49(3): 231–236. https://doi.org/10.1016/j.ifacol.2016.07.039

Rosić, M.; Pešić, D.; Kukić, D.; Antić, B.; Božović, M. 2017. Method for selection of optimal road safety composite index with examples from DEA and TOPSIS method, *Accident Analysis* & *Prevention* 98: 277–286. https://doi.org/10.1016/j.aap.2016.10.007

Rostamzadeh, R.; Akbarian, O.; Banaitis, A.; Soltani, Z. 2021. Application of DEA in benchmarking: a systematic literature review from 2003–2020, *Technological and Economic Development of Economy* 27(1): 175–222. https://doi.org/10.3846/tede.2021.13406

Sexton, T. R.; Silkman, R. H.; Hogan, A. J. 1986. Data envelopment analysis: critique and extensions, *New Directions for Program Evaluation* 1986(32): 73–105. https://doi.org/10.1002/ev.1441

Shen, Y.; Hermans, E.; Bao, Q.; Brijs, T.; Wets, G. 2013. Road safety development in Europe: a decade of changes (2001–2010), *Accident Analysis* & *Prevention* 60: 85–94. https://doi.org/10.1016/j.aap.2013.08.013

Shen, Y.; Hermans, E.; Bao, Q.; Brijs, T.; Wets, G. 2015. Serious injuries: an additional indicator to fatalities for road safety benchmarking, *Traffic Injury Prevention* 16(3): 246–253. https://doi.org/10.1080/15389588.2014.930831

Shen, Y.; Hermans, E.; Brijs, T.; Wets, G.; Vanhoof, K. 2012. Road safety risk evaluation and target setting using data envelopment analysis and its extensions, *Accident Analysis* & *Prevention* 48: 430–441. https://doi.org/10.1016/j.aap.2012.02.020

Shen, Y.; Hermans, E.; Ruan, D.; Wets, G.; Brijs, T.; Vanhoof, K. 2011. A generalized multiple layer data envelopment analysis model for hierarchical structure assessment: a case study in road safety performance evaluation, *Expert Systems with Applications* 38(12): 15262–15272. https://doi.org/10.1016/j.eswa.2011.05.073

Si, Q.; Ma, Z. 2019. DEA cross-efficiency ranking method based on grey correlation degree and relative entropy, *Entropy* 21(10): 966. https://doi.org/10.3390/e21100966

Stanković, M.; Stević, Ž.; Das, D. K.; Subotić, M.; Pamučar, D. 2020. A new fuzzy MARCOS method for road traffic risk analysis, *Mathematics* 8(3): 457. https://doi.org/10.3390/math8030457

Su, C.-H.; Lu, T. 2019. An entropy-based cross-efficiency under variable returns to scale, *Entropy* 21(12): 1205. https://doi.org/10.3390/e21121205

Tsaples, G.; Papathanasiou, J. 2021. Data envelopment analysis and the concept of sustainability: a review and analysis of the literature, *Renewable and Sustainable Energy Reviews* 138: 110664. https://doi.org/10.1016/j.rser.2020.110664

Tsolas, I. E. 2019. Utility exchange traded fund performance evaluation. a comparative approach using grey relational analysis and data envelopment analysis modelling, *International Journal of Financial Studies* 7(4): 67. https://doi.org/10.3390/ijfs7040067

Wang, Y.-M.; Chin, K.-S. 2011. The use of OWA operator weights for cross-efficiency aggregation, *Omega* 39(5): 493–503. https://doi.org/10.1016/j.omega.2010.10.007

Wang, Z.-X.; Yao, P.-Y. 2018. Grey relational analysis of economic policy uncertainty in selected European union countries, *Economic Computation and Economic Cybernetics Studies and Research* 52(2): 251–265. https://doi.org/10.24818/18423264/52.2.18.15

Xiong, X.; Chen, L.; Liang, J. 2018. Analysis of roadway traffic accidents based on rough sets and Bayesian networks, *Promet – Traffic* & *Transportation* 30(1): 71–81. Available from Internet: https://traffic2.fpz.hr/index.php/promtt/article/view/2713

Yang, Y.; Liu, S.; Xie, N. 2019. Uncertainty and grey data analytics, *Marine Economics and Management* 2(2): 73–86. https://doi.org/10.1108/maem-08-2019-0006

Yushimito, W. F.; Alves, P. N.; Canessa, E.; De Mateo, F. 2018. Relating efficiency with service compliance indices in public transportation using slack-based measure data envelopment analysis and shadow prices, *Promet – Traffic* & *Transportation* 30(6): 661–670. Available from Internet: https://traffic2.fpz.hr/index.php/promtt/article/view/2767

Zakowska, I.; Godycki-Cwirko, M. 2020. Data envelopment analysis applications in primary health care: a systematic review, *Family Practice* 37(2): 147–153. https://doi.org/10.1093/fampra/cmz057

Zhu, J.; Hipel, K. W. 2012. Multiple stages grey target decision making method with incomplete weight based on multi-granularity linguistic label, *Information Sciences* 212: 15–32. https://doi.org/10.1016/j.ins.2012.05.011