

Research Article

Application of Fuzzy Linguistic Rating and Entropy-Based GRA to Address Uncertainty in Safety Performance Index Construction

Vladimir Pajković ¹, Mirjana Grdinić-Rakonjac ¹ and Boris Antić ²

¹Department of Road Traffic, Faculty of Mechanical Engineering, University of Montenegro, Podgorica 81000, Montenegro

²Department of Road Safety, Faculty of Transport and Traffic Engineering, University of Belgrade, Belgrade 11000, Serbia

Correspondence should be addressed to Mirjana Grdinić-Rakonjac; grdnicm@ucg.ac.me

Received 26 August 2021; Revised 15 January 2022; Accepted 29 January 2022; Published 1 March 2022

Academic Editor: Richard S. Tay

Copyright © 2022 Vladimir Pajković et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As the need to assess the level of road safety grows, there is a noticeable tendency of experts to use one overall composite index that contains information on a number of safety performance indicators (SPIs). Indicators commonly used in road safety assessment are numerical, and their natural uncertainty and vagueness are often overlooked. However, there are also SPIs that are rather linguistic, such as data on driver behavior, which are most often collected through questionnaires and are considered qualitative, imprecise, and fuzzy. Together with inappropriate selection of weighting and aggregation methods, such data can be a source of uncertainty and can lead to unreliable results and erroneous conclusions. In this regard, the present study provides a systematic and efficient hybrid method that integrates three different procedures to deal with unavoidable uncertainty in each step of index construction. The application of fuzzy linguistic rating grasps insight into the ambiguity that is intrinsic in drivers' self-assessment. Entropy describes each observed behavior by quantifying the disorder of a system. Grey relational analysis aggregates behavioral indicators into a composite index, doubting their sufficiency and completeness. A case study of Montenegro has been provided to demonstrate the practical applicability of the proposed method in safety assessment under uncertainty. Results abstracted not wearing the seatbelt as the most common negative behavior among drivers in Montenegro, followed by using the telephone while driving, speeding, and driving under the influence of alcohol. In addition, municipalities are ranked according to the level of road safety.

1. Introduction

The aggregation of different variables into one safety performance index is a popular concept in evaluating road safety and in comparing the performance of territories. Researchers have established a comprehensive set of safety performance indicators (SPIs), namely composite indicators, that takes into account both direct and indirect influences of road safety risk factors. Hereupon, a large number of SPI combinations were developed to create an overall composite index, for example [1–3].

Indicators commonly used in road safety assessment are numerical and their natural uncertainty and vagueness are often overlooked. The data collected may sometimes be the result of a measurement error or the use of proxy data,

whose reliability is, therefore, questionable [4]. Furthermore, SPIs that describe drivers' attitudes and behavior are usually collected by measuring with physical devices and from questionnaires [5–7]. The indicator value, which is usually the mean value of all the respondents' answers, cannot be considered completely accurate [8]. Not only by arbitrarily and carelessly recording opinions and attitudes but also by not understanding the questionnaire (or some part of it), respondents can cause uncertainty and vagueness of the collected data [9]. Respected that a road safety score is as reliable as the data on which it is constructed and that its quality should improve in parallel with the improvement of the quality of the information [10], the question arises about the composite index constructed from such type of data. In addition, the selection of weighting and aggregation

techniques in the index construction process can be a source of the unreliability of the obtained results. Therefore, ignoring uncertainty as the inherent feature of the mentioned kind of data will lead to unreliable results, and it is necessary to consider uncertain parameters to preserve the robustness of the safety performance index.

Some authors have suggested different methods for handling imprecise data, and its models are singled out as superior and common to capture uncertainty. The existing models for road safety evaluation commonly consider the data uncertainty in one of the steps: data modeling [11], weighting [12–14], or aggregation [15, 16], and they usually do so only when the subjective opinion of road safety experts is involved [17, 18]. The main aim of this paper is to propose a hybrid method for constructing a road safety performance index that will consider the vagueness and uncertain nature of SPIs related to driver behavior, suggesting the integration of uncertainty-solving techniques in each step. Such a method is needed because most of the countries still do not have a reliable database on road safety, do not have enough accurate data on traffic safety, and data on driver behavior are most often collected using questionnaires. The main contribution of the proposed approach is reflected as follows: both qualitative and quantitative input data can be considered and aggregated together into one composite index. The weighting of all indicators can be obtained without the implementation of any subjective information from the outside, reflecting at the same time the uncertainty of selected indicators. The proposed aggregation technique allows handling uncertainty for system evaluation when the decision makers are unsure about the sufficiency and completeness of the data at hand by providing a set of robust road safety performance indexes for each territory that allow full ranking. Robust and logical results verify the possibility of the application of the proposed method in many other fields besides road safety.

The rest of the paper is organized as follows: in the next section, the feature of SPIs will be discussed and ways to address their natural uncertainty will be outlined. In Section 3, the case study will be presented, followed by the results in Section 4. In Sections 5 and 6, the research will be discussed and concluded.

2. Uncertain Nature of Safety Performance Indicators

It was confirmed that every safety analysis has some degree of uncertainty [10]. Indicators that are usually used to describe the state of road safety rely on empirical data, such as the number of fatalities and injured, the number of accidents, the motorization level, [14, 19], etc. Although crisp, because of real-life application, there are numerous reasons for the uncertainty, imprecision, or greyness of these data values. Measurement error and proxy data often cause relevant data incurrence. For example, recorded SPIs, such as the number of drivers on the roads speeding or driving under the influence of alcohol, are not fully known because it is impossible to get accurate information. SPIs based on traffic violations that are used as indicators by some

researchers, [20, 21], are also a reason for the vagueness of the final score since this type of data can represent the approximate situation on roads but not the precise state of safety and drivers' behavior. Likewise, Lloyd and Forster [22] stated that risk exposures (especially accident risk, which is defined as the rate of accidents per vehicle kilometer traveled) represent uncertain information. Hence, the reliability of the performance index derived from this kind of data is questionable. However, none of the mentioned studies took into account the uncertainty of the data in the matter when calculating the composite index, despite the proven uncertain nature of these data.

Furthermore, data related to self-reported behavior, which are often used as SPIs in index construction, represent a qualitative measure of opinion about one's own behavior. When collecting data on behavior and perception (usually through questionnaires), the main focus is on their final presentation as, in many cases, the respondents are not able to clearly express their judgments. Besides road users, decision makers are sometimes involved in trying to express their opinion and preference on particular road safety domains. Since the precision of these subjective measures is unlikely, it is assumed that all final values derived from self-reported data and perception data are imprecise [23] and fuzzy [24], and identifying and addressing uncertainties is essential to manipulating this data. The intention to handle them can be found in the literature.

The simplest way to cope with uncertain data is averaging, which implies statistically abstracting the probability distribution (by means and variance). Tsang et al. [25] applied decision trees to handle data sets with uncertain values and showed that the results are more accurate than those using an average value. Xiong et al. [11] analyzed roadway traffic accidents based on rough sets (for data) and Bayesian networks (for aggregation). Chen et al. [12] performed a road safety risk evaluation applying entropy weighting. Entropy weighting is also used in [14] to assess urban road safety and to rank the provinces accordingly. Ayati et al. [26] applied the evidential reasoning (ER) approach to consider the subjective state of evaluation of the roadside hazard severity and to assess the situation more realistically. Rassafi et al. [10] collected the opinions of road users and experts for the purpose of assigning weights in the process of road safety assessment. They also used the ER approach in addition to Dempster–Shafer theory to handle the uncertainty of data. Of the other multicriteria decision-making methods, fuzzy-oriented analytic hierarchy process [18, 27] and data envelopment analysis [28, 29] were mainly used. It is proven that despite the differences, all existing models for addressing uncertainties are related to one another [30].

The widely used fuzzy theory [31] is an appropriate method to process imprecise and inaccurate data, notably when they are described in human language. The same could be said for the grey theory [32]. Both fuzzy and grey theories deal with the uncertainty of human perception in a manner that creates a mechanism of similar reasoning. The application of these theories for addressing uncertainty will be described in the next sections.

2.1. Fuzzy Linguistic Rating. When applying questionnaires to collect road safety behavioral data, the authors are usually conflicted about whether to use parametric or nonparametric scale procedures since it is not fully clear if the data are interval or ordinal. Questionnaires that are usually employed to collect perceptions, ratings, or judgments on many subjects in various fields rely on the most famous and widely used Likert scale [33, 34]. Likert scale traditionally applies linguistic variables that are encoded by means of ranking (in most cases, from 1 to 5–7, e.g., from the least “strongly disagree” to the largest “strongly agree”). This way of measuring opinions is very popular since it is easily conducted. Based on the answers, the final data represent the average value of even-spaced ranks with a given possibility of applying some statistical tools. However, this type of scale has some drawbacks. For instance, some of the information may be lost because of the imprecision of selecting one single value of the given 5–7 options on the rating scale, while the space between each rank cannot be equidistant. Furthermore, the transition between the ranks may be differently perceived based on different respondents, and most importantly, the interpretation of results is not very reliable because applying statistical tools in most cases is limited [33]. In that matter, some authors turned to the usage of the fuzzy theory that refers to addressing the uncertainty and imprecision of data and that relies on possibility and neglects probability [35].

As a part of the fuzzy theory, fuzzy numbers have already been recognized as a tool in psychometric studies when imprecision and subjectivity need to be captured. There are two ways of dealing with these issues. One way is to introduce an alternate scale, question formats, etc., and the other is to develop new methods to analyze this type of collected data. Recently, applying the fuzzy set theory to capture the subjectivity and imprecision of each collected answer has become widespread. In addition, fuzzy numbers are employed to grasp an insight into the ambiguity that is intrinsic to human assessment and rating [9]. In the literature, this method is known as the fuzzy linguistic approach and has been used by many authors in different fields.

The fuzzy linguistic rating scale is most commonly used for obtaining the criteria weights in multiattribute decision problems. Bao et al. [36] used triangular fuzzy numbers to calculate the criteria weights for composite index construction. Mousavi et al. [37] used trapezoidal fuzzy numbers to rate the alternative and the criteria in the manufacturing system. Tseng and Chiu [38] calculated the weights of the green supply chain based on fuzzy linguistic

preferences. Awasthi and Kannan [39] applied the triangular fuzzy numbers to derive criteria and alternative weights for the purpose to develop suppliers’ environmental performance. A triangular fuzzy rating scale was used by Stanković et al. [40] and Rostamzadeh et al. [41] to evaluate solutions for road sections and to calculate the significance of the selected transport and logistic problems, respectively. Wersi Qazvini et al. [42] applied the fuzzy TOPSIS approach for the aggregation of alternatives in the process of identifying and analyzing the black spots of suburban areas, while Memiş et al. [17] evaluated the criteria using the fuzzy pivot pairwise relative criteria importance assessment approach to determine and rank the road transportation risk factors.

The above-mentioned studies validate the fuzzy linguistic rating scale as a tool for capturing the imprecision and uncertain nature of the Likert scale. As can be seen, the triangle and trapezoid are the most commonly used shapes of a fuzzy number, representing the compromise between narration and calculation. Unlike the previous practice of using fuzzy numbers to describe decision makers’ opinions, and based on them, to calculate weights of indicators, in this paper, they are employed to describe the driver’s self-reported behavior and accordingly calculate the values of indicators.

The result of the fuzzy approach counts heavily on the degree that fuzzy numbers correctly represent qualitative data, i.e., on assigning the appropriate membership functions for fuzzy numbers. In this paper, the fuzzy rating of each input indicator is described by linguistic expressions. Calculations were done by transferring these variables into trapezoidal fuzzy numbers (i.e., a_1 , a_2 , a_3 , and a_4) with the following membership function:

$$\begin{aligned}\mu_{a(x)} &= \frac{x - a_1}{a_2 - a_1}; x \in [a_1, a_2], \\ \mu_{a(x)} &= 1; x \in [a_2, a_3], \\ \mu_{a(x)} &= \frac{a_4 - x}{a_4 - a_3}; x \in [a_3, a_4], \\ \mu_{a(x)} &= 0; \text{ otherwise}\end{aligned}\quad (1)$$

The fuzzy rating scale of the frequency of a particular driving behavior of each respondent is described with one of the six linguistic expressions given in Table 1, Defuzzified value represents final (crisp) value that will serve as input data in the behaviour index construction.

$$i_{ij} = \text{Defuzz}(x_{ij}) = \frac{(1/3)(x_{ij4} - x_{ij3})^2 + (1/3)(x_{ij2} - x_{ij1})^2 - x_{ij1}x_{ij2} + x_{ij3}x_{ij4}}{-x_{ij1} - x_{ij2} + x_{ij3} + x_{ij4}} \quad (2)$$

In that manner, a fuzzy rating scale was used here to model a road user’s self-reported behavior, similar to

[38, 43]. Furthermore, on such data, two more methodologies will be applied for weighting and aggregating, with the

TABLE 1: Transformation of linguistic variables to fuzzy numbers.

Linguistic variable	Fuzzy number
Never	(0.0, 0.0, 0.1, 0.2)
Rarely	(0.1, 0.2, 0.3, 0.4)
Sometimes	(0.3, 0.4, 0.4, 0.5)
Often	(0.4, 0.5, 0.6, 0.7)
Very often	(0.6, 0.7, 0.8, 0.9)
Always	(0.8, 0.9, 1.0, 1.0)

main aim to model all existing uncertainties and to construct one reliable and robust overall road safety behavior index under an uncertain environment.

2.2. Weighting and Aggregating Methods. Combining SPIs into a complex performance index includes assessing the relative importance of each indicator (assigning weights) and aggregating indicators.

As a weighting technique based on statistical parameters that measure the amount of information in a variable, Shannon's entropy method can assess the importance of the indicators by reflecting the unevenness of the dataset. Hence, it is suitable for reflecting the uncertainty of behavioral indicators. Hence, in the context of traffic safety, it has recently been widely used to determine the criteria weights in the form of a common set of weights [12, 14, 44].

In addition, the grey theory is a useful model for the analysis of uncertain systems with partially known and partially unknown information in many fields. Grey relational analysis (GRA), as a part of a grey theory, represents a normalization-based technique, which implies the positive values of the data sequence generated by translating the performance of all alternatives into series, and unlike other methods for addressing uncertainties, it works well on small sample size. A combination of fuzzy theory and GRA was used by Ma et al. [45] for the purpose of creating three different sets of road safety indicators related to experts' attitudes: one related to highways, one related to urban roads, and one related to regional roads. Hu et al. [46] used GRA in addition to fuzzy and data envelopment analysis to develop public transport network evaluation, and both Liu et al. [47] and Grdinić-Rakonjac et al. [48] used GRA to calculate the weights of road safety indicators. In this paper, GRA was applied to the aggregation of weighted indicators and the index construction, as shown in Figure 1.

With this novel hybrid methodology (entropy-based weighting + GRA aggregating), it is possible to create a reliable composite safety performance index of the territory and additionally establish a more rational ranking among the territories under evaluation. The steps of entropy and grey relational analysis are summarized in Table 2.

3. Case Study

The proposed approach is tested for a set of 21 Montenegrin municipalities. Since one of the goals of this research was to identify the most influential road safety risk behavior, the driver is chosen to represent the unit of the analysis rather

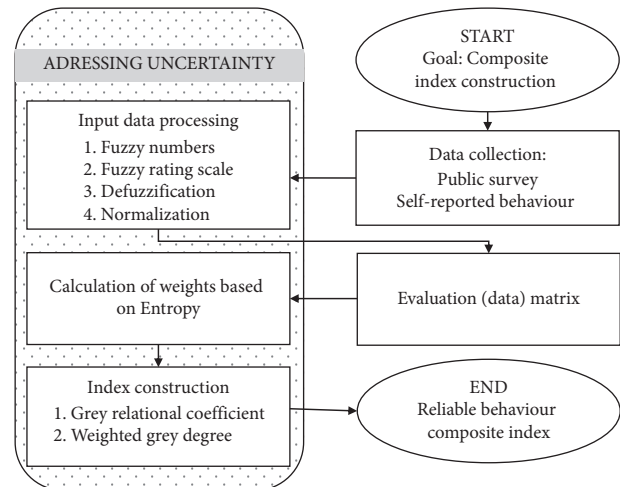


FIGURE 1: Flow chart of the proposed methodology.

than any other users involved in the road traffic. The following unwanted behaviors were investigated: driving above the speed limit, driving under the influence of alcohol, using the telephone while driving, and not wearing the seatbelt. Data were collected by conducting face-to-face and online questionnaire surveys among 1309 drivers with a priori determined sample size for each municipality (3% error bound and 95% confidence limits). Descriptive statistics are given in Table 3.

The questionnaire was compiled in accordance with the ESRA and SARTRE methodology, which means that each respondent describes the frequency of certain risk behaviors while driving using one of the values of the language variables: never, rarely, sometimes, often, very often, and always. All questions are divided into those related to main roads, regional roads, and local/urban roads. Finally, drivers answered up to 25 questions grouped into four main domains (V1–V9, A1–A3, T1–T4, and S1–S9) and presented in Table 4. Their answers are given in Table 5 in percentage.

4. Results

The fuzzy theory refers to addressing the uncertainty and imprecision of data. Hence, in this paper, all self-reported behaviors were processed and converted into numerical values using trapezoidal fuzzy numbers and their operations. The data are organized in such a way that higher values represent positive behavior. The final fuzzy number (Table 6) for the particular risk behavior (for example, speeding in the municipality AN) was derived as the weighted sum of all calculated fuzzy numbers for each question. Internal weights (see Figure 2) were determined by experts relying on the road safety figures (number of accidents and fatalities on different types of roads). Defuzzification was conducted via equation (2), and the final values were obtained (Table 7). Not wearing the seatbelt received the lowest average score (0.568), abstracting this negative behavior as the most common among drivers in Montenegro, followed by using the telephone while driving, speeding, and driving under the influence of alcohol (0.593, 0.650, and 0.881, respectively).

TABLE 2: Steps of entropy and grey relational analysis.

Step	Entropy	Grey relational analysis
I	$p_{ij} = x_{ij} / \sum_{i=1}^m x_{ij}$	$g_j(k) = (x_j(k) - \min_{j=1}^n x_j(k)) / (\max_{j=1}^n x_j(k) - \min_{j=1}^n x_j(k))$
II	$E_j = -(1/\ln(m)) \sum_{i=1}^m p_{ij} \ln p_{ij}$	$\zeta_{0i}(j) = (\min_{i=1}^n \min_{j=1}^m g_i(j) - g_0(j) + \rho \max_{i=1}^n \max_{j=1}^m g_i(j) - g_0(j)) / (g_i(j) - g_0(j) + \rho \max_{i=1}^n \max_{j=1}^m g_i(j) - g_0(j))$
III	$d_j = 1 - E_j$	$\Gamma_{0i}(j) = (1/n) \sum_{j=1}^m \zeta_{0i}(j)$
IV	$W_j = d_j / \sum_{j=1}^n d_j$	$\Gamma_{0j}(j) = (1/n) \sum_{k=1}^l \zeta_{0j}(j) W(j)$

TABLE 3: Descriptive statistics of the final sample.

Variable	Total	
	No.	%
Gender	Males	707 (54.01)
	Females	602 (45.99)
Licensure	≤2 years	158 (12.06)
	2–5 years	217 (16.61)
	>5 years	934 (71.32)
Driving frequency	Daily	998 (76.24)
	1–3 times per week	195 (14.90)
	1–3 times per week	55 (4.20)
	Less than 1 time per week	61 (4.66)
		1309 (100)

Following the steps described in Section 2, the entropy-based weighting technique was introduced, and it was found that not wearing a seatbelt had the greatest influence on the final score with an assigned value of 0.528. The lowest assigned weight of 0.046 was received by the behavior related to driving under the influence of alcohol (first row in Table 8).

Furthermore, these weights were integrated with grey relational coefficients, and the weighted grey degree $\Gamma_{0j}(j)$ was obtained, representing the final behavior FEGRA index. Results are presented in Table 8. The municipality of Tivat (TV) received the largest behavior index (0.778), indicating that the local drivers have relatively better behavior than drivers in other territories. It is followed by municipality of Kotor (KO), Bar (BR), Šavnik (ŠA), and Mojkovac (MK). These scores are not a surprise given that those municipalities are all small ones, with mostly low motorization levels and no main roads on their territories. In contrast, the municipality of Budva (BD) was assigned a relatively low score (0.371) and was identified as the worst-performing municipality, along with the municipality of Žabljak (ŽB). Both of these municipalities represent tourist centers. Hence, an increased number of vehicles and population on their territories may be the cause of low performance.

5. Discussion

Having in mind that the construction of a composite index implies (more or less subjective) methodological choices in several consecutive steps, the robustness analysis of the proposed methodology, as a kind of tool for quality assurance [49], was performed. The influence of data modeling methods, weighting, and aggregation schemes on final outcomes (i.e., a shift in the rank of the entire set of

municipalities) was examined. In addition to this uncertainty analysis, the Pearson product-moment correlation ratio was chosen as a variance-based measure that accounts for linear dependence between the input and output data, and it is exactly correspondent to the main effect index or the first-order sensitivity index [50, 51].

In the first step, the impact of the data modeling method will be examined. As it was said earlier, the FEGRA index is constructed with the data collected from the questionnaire. A common practice in data modeling, when opinion and attitudes are involved, is the usage of the Likert scale. It is usually done by averaging the collected answers. Firstly, the survey answers were observed as ordinal with “1,” presenting unwanted behavior, and “6,” presenting positive behavior, and the final crisp score for the selected behavior was calculated as the mean of all collected answers by respecting the previously defined internal weights. Those values are then normalized (see Table 9). However, the uncertainty of the drivers’ subjective measure of their own behavior is not addressed in that case. The comparative influence of the applied fuzzy numbers for resolving the uncertainty of input data is presented in Figure 3. The calculated rank-order coefficient of 0.69 shows a relatively low correlation between rankings, meaning that the usage of fuzzy numbers had a significant effect on the final performance index. Figure 3 illuminates that the largest rank variation is recorded for two top-ranked territories (rank shift is 13 and 8 places, respectively). Besides this, the data modeling technique used affects principally the bottom-ranked territories.

In the second step, the influence of the weighting method is considered. The weights attached to each indicator are usually chosen to reflect the importance of that indicator with respect to the concept being measured [50]. An ordinary initial approach is to take that each input variable

TABLE 4: The survey questionnaire.

Survey questions	Mark
<i>Driving above the speed limit</i>	
How often do you drive over the speed limit?	
[On main roads]	V1
[On regional roads]	V2
[On local/urban roads]	V3
Do you exceed the prescribed speed on roads with higher permitted speeds (60–80 km/h)?	
[By more than 30 km/h]	V4
[By less than 30 km/h]	V5
Do you exceed the prescribed speed on roads with lower permitted speeds (50–60 km/h)?	
[By more than 20 km/h]	V6
[By less than 20 km/h]	V7
Do you exceed the prescribed speed on local/urban roads?	
[By more than 10 km/h]	V8
[By less than 10 km/h]	V9
<i>Driving under the influence of alcohol</i>	
Do you drive under the influence of alcohol?	
[On main roads]	A1
[On regional roads]	A2
[On local/urban roads]	A3
<i>Wearing the seatbelt</i>	
How often do you wear your seatbelt while driving?	
[On main roads]	P1
[On regional roads]	P2
[On local/urban roads]	P3
How often your front passenger is wearing a seatbelt while driving?	
[On main roads]	P4
[On regional roads]	P5
[On local/urban roads]	P6
Do you use the child resistant system for children under the age of 5?	
[On main roads]	P7
[On regional roads]	P8
[On local/urban roads]	P9
<i>Using mobile phone while driving</i>	
Do you use a hands-free device to talk to a mobile phone while driving?	
[On main or regional roads]	T1
[On local/urban roads]	T2
Do you use a mobile phone while driving to write messages and visit social networks?	
[On main or regional roads]	T3
[On local/urban roads]	T4

contributes equally. This equal weighting is the most common scheme appearing in composite index construction [52]. Also, lately, the very popular method for weighting is the data envelopment analysis (DEA) since, as a data-driven technique, it does not rely on a subjective measure from outside. Here, a comparison of these two methods with ranking according to the proposed entropy weighting is performed. The DEA model from [43] is used, and separate weights for each territory are obtained. Figure 4 (left graph) compares these rankings, and it can be seen that addressing data uncertainty using the entropy weighting scheme has a positive impact on the ranking of eight and twelve territories compared to equal and DEA weighting, respectively, whilst a negative effect on the rank order is recorded in twelve and six territories, respectively. The higher correlation coefficient (0.72 and 0.70 for equal weighting and DEA, respectively) indicates that the weighting method has less (but not negligible) impact than data modeling.

In terms of aggregation techniques, the additive (simple linear) approach and the multicriteria decision-making (MCDM) approach were chosen for comparison. Additive aggregation, as the most commonly used in composite indicators, allows the assessment of the marginal contribution of each variable separately and entails full compensability. However, when assigned weights are considered to be “importance coefficients” (weights representing a measure of importance), non-compensatory aggregation should be used for composite index construction [52]. In that sense, the MCDM-TOPSIS method was taken, considering the preferential order of indicators by similarity with the ideal solution. Figure 4 (right graph) compares the ranking obtained by both additive and MCDM aggregation with the ranking obtained by FEGRA when the grey relational analysis is employed. The relatively high correlation coefficients between rankings under the study show a somewhat

TABLE 5: Collected answers (in %).

Speeding	Never	Rarely	Sometimes	Often	Very often	Always
V1	20.55	28.27	29.34	12.22	5.19	4.43
V2	31.26	30.20	23.24	9.01	3.10	3.18
V3	44.00	29.33	16.95	5.40	1.82	2.51
V4	44.24	27.73	16.29	7.12	2.73	1.89
V5	24.28	28.07	20.94	13.81	5.92	6.98
V6	36.66	28.12	21.69	8.92	2.57	2.04
V7	23.54	25.06	25.74	13.25	6.06	6.36
V8	41.39	26.07	20.88	6.78	2.67	2.21
V9	30.57	26.51	19.31	12.11	4.83	6.67
Alcohol						
A1	75.28	20.67	4.04	0.00	0.00	0.00
A2	82.76	11.09	4.27	0.97	0.52	0.37
A3	76.39	12.89	6.52	2.70	0.90	0.60
Telephone						
T1	67.30	10.57	8.69	4.76	2.42	6.27
T2	70.64	10.34	7.70	4.08	2.42	4.83
T3	60.35	18.58	11.03	5.97	2.11	1.96
T4	58.46	15.86	10.95	6.95	4.15	3.63
Seatbelt						
S1	6.44	6.44	10.71	9.74	7.27	59.40
S2	9.51	7.86	13.47	9.58	7.04	52.54
S3	15.45	8.25	12.00	9.53	7.80	46.96
S4	8.86	7.96	13.66	9.68	12.69	47.15
S5	9.48	9.93	15.82	11.19	13.28	40.30
S6	13.89	10.81	17.72	9.91	12.31	35.36
S7	15.75	4.94	11.28	7.52	7.87	52.64
S8	16.37	6.01	10.84	7.30	9.42	50.06
S9	15.58	7.08	12.16	8.38	9.45	47.34

TABLE 6: Assigned fuzzy numbers of self-reported behavior.

	Speeding	Alcohol	Telephone	Seatbelt
AN	(0.558, 0.611, 0.736, 0.793)	(0.712, 0.811, 0.909, 0.939)	(0.495, 0.566, 0.660, 0.711)	(0.420, 0.494, 0.571, 0.634)
BA	(0.575, 0.629, 0.756, 0.811)	(0.708, 0.806, 0.904, 0.935)	(0.499, 0.572, 0.667, 0.720)	(0.398, 0.470, 0.539, 0.594)
BD	(0.418, 0.472, 0.577, 0.649)	(0.695, 0.795, 0.891, 0.918)	(0.399, 0.464, 0.558, 0.621)	(0.373, 0.437, 0.509, 0.555)
BP	(0.552, 0.606, 0.734, 0.796)	(0.764, 0.864, 0.964, 0.976)	(0.505, 0.584, 0.678, 0.732)	(0.488, 0.572, 0.648, 0.697)
BR	(0.580, 0.632, 0.762, 0.818)	(0.765, 0.865, 0.965, 0.978)	(0.512, 0.590, 0.685, 0.737)	(0.533, 0.617, 0.694, 0.735)
CT	(0.557, 0.613, 0.745, 0.801)	(0.669, 0.767, 0.864, 0.896)	(0.450, 0.527, 0.620, 0.683)	(0.392, 0.464, 0.546, 0.605)
DG	(0.603, 0.658, 0.788, 0.842)	(0.779, 0.879, 0.979, 0.990)	(0.473, 0.547, 0.638, 0.693)	(0.417, 0.490, 0.558, 0.606)
HN	(0.404, 0.454, 0.563, 0.636)	(0.680, 0.780, 0.870, 0.904)	(0.458, 0.534, 0.627, 0.690)	(0.492, 0.568, 0.645, 0.682)
KL	(0.581, 0.635, 0.765, 0.826)	(0.737, 0.837, 0.934, 0.956)	(0.452, 0.530, 0.620, 0.686)	(0.495, 0.574, 0.651, 0.689)
KO	(0.445, 0.496, 0.611, 0.680)	(0.672, 0.771, 0.866, 0.901)	(0.494, 0.569, 0.664, 0.717)	(0.620, 0.706, 0.789, 0.805)
MK	(0.548, 0.599, 0.719, 0.776)	(0.774, 0.874, 0.974, 0.987)	(0.550, 0.636, 0.732, 0.781)	(0.427, 0.506, 0.583, 0.643)
NK	(0.545, 0.602, 0.730, 0.797)	(0.735, 0.834, 0.933, 0.955)	(0.459, 0.534, 0.630, 0.691)	(0.528, 0.615, 0.696, 0.740)
PG	(0.551, 0.607, 0.737, 0.804)	(0.754, 0.853, 0.952, 0.967)	(0.505, 0.585, 0.677, 0.733)	(0.474, 0.555, 0.635, 0.685)
PL	(0.478, 0.538, 0.649, 0.724)	(0.770, 0.870, 0.968, 0.979)	(0.331, 0.397, 0.486, 0.568)	(0.413, 0.487, 0.554, 0.601)
PV	(0.519, 0.575, 0.697, 0.764)	(0.747, 0.847, 0.943, 0.958)	(0.522, 0.597, 0.695, 0.739)	(0.375, 0.444, 0.506, 0.558)
PŽ	(0.593, 0.645, 0.777, 0.825)	(0.769, 0.869, 0.969, 0.980)	(0.431, 0.505, 0.594, 0.657)	(0.379, 0.451, 0.513, 0.562)
RO	(0.527, 0.573, 0.696, 0.753)	(0.762, 0.862, 0.958, 0.968)	(0.399, 0.474, 0.556, 0.625)	(0.334, 0.396, 0.466, 0.525)
ŠA	(0.607, 0.660, 0.796, 0.847)	(0.766, 0.866, 0.966, 0.975)	(0.519, 0.600, 0.695, 0.744)	(0.452, 0.528, 0.607, 0.652)
TV	(0.532, 0.584, 0.715, 0.776)	(0.654, 0.750, 0.843, 0.870)	(0.446, 0.522, 0.613, 0.676)	(0.639, 0.728, 0.818, 0.839)
UL	(0.491, 0.551, 0.664, 0.740)	(0.772, 0.872, 0.970, 0.979)	(0.412, 0.485, 0.574, 0.640)	(0.425, 0.498, 0.568, 0.612)
ŽB	(0.405, 0.453, 0.561, 0.641)	(0.747, 0.847, 0.943, 0.954)	(0.447, 0.516, 0.612, 0.670)	(0.369, 0.421, 0.487, 0.521)

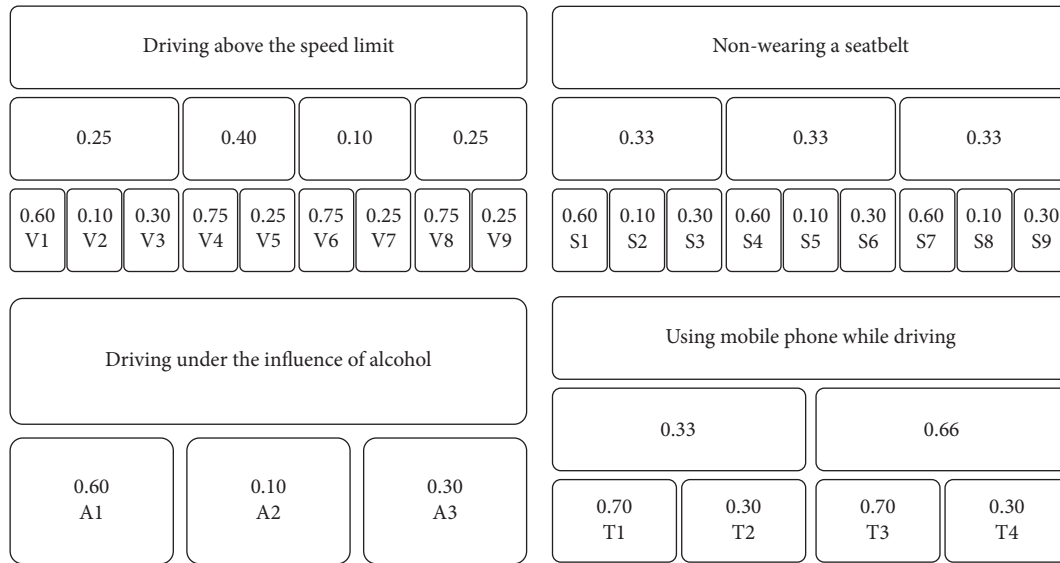


FIGURE 2: Internal weights of risk behavior.

TABLE 7: Defuzzificated and normalized self-reported behavior values.

	Defuzzificated				Normalized			
	Speeding	Alcohol	Telephone	Seatbelt	Speeding	Alcohol	Telephone	Seatbelt
AN	0.680	0.860	0.618	0.542	0.753	0.493	0.700	0.306
BA	0.698	0.855	0.625	0.512	0.839	0.455	0.729	0.218
BD	0.535	0.843	0.520	0.478	0.072	0.361	0.278	0.116
BP	0.677	0.911	0.637	0.616	0.741	0.882	0.780	0.531
BR	0.703	0.912	0.643	0.660	0.860	0.891	0.805	0.663
CT	0.685	0.816	0.582	0.513	0.775	0.153	0.543	0.220
DG	0.728	0.926	0.599	0.530	0.980	1.000	0.617	0.273
HN	0.520	0.827	0.588	0.609	0.000	0.236	0.572	0.511
KL	0.707	0.885	0.584	0.616	0.881	0.682	0.554	0.531
KO	0.563	0.820	0.622	0.746	0.204	0.182	0.716	0.922
MK	0.666	0.921	0.688	0.553	0.687	0.961	1.000	0.342
NK	0.674	0.882	0.590	0.661	0.727	0.663	0.577	0.665
PG	0.681	0.900	0.637	0.601	0.758	0.799	0.782	0.486
PL	0.605	0.915	0.455	0.527	0.399	0.920	0.000	0.262
PV	0.645	0.893	0.649	0.483	0.588	0.745	0.831	0.130
PŽ	0.715	0.915	0.558	0.490	0.917	0.918	0.441	0.150
RO	0.642	0.907	0.525	0.440	0.573	0.852	0.302	0.000
ŠA	0.732	0.912	0.652	0.573	1.000	0.892	0.844	0.399
TV	0.657	0.797	0.576	0.772	0.644	0.000	0.519	1.000
UL	0.619	0.917	0.539	0.538	0.466	0.934	0.360	0.297
ŽB	0.520	0.892	0.571	0.457	0.001	0.738	0.497	0.052

negligible influence on the final score, with the total absolute average shift in the rank of 1.05 places when GRA is replaced with additive aggregation and 1.5 places when aggregation is conducted with the MCDM technique (Table 10). In Table 10, rank shifts are given (individually by municipalities and total) along with the measures of linear dependence–Pearson product-moment correlations for all considered alternatives.

Finally, the ranking according to the complete FEGRA method was compared with the one performed by the selected combination of steps in the development of the composite index: the Likert scale (data modeling)–equal weight (weighting)–additive approach (aggregation). This comparison is illustrated in Figure 5 with error bars representing the deviation of FEGRA rank by ± 3 places. From the scatter plot, it can be seen that larger deviations

TABLE 8: Final behavior index and rank.

Entropy weights	Grey degree				FEGRA-behaviour index and (rank)
	Speeding	Alcohol	Telephone	Seatbelt	
	0.237	0.046	0.189	0.528	
AN	0.669	0.496	0.625	0.419	0.521 (13)
BA	0.756	0.479	0.648	0.390	0.530 (11)
BD	0.350	0.439	0.409	0.361	0.371 (21)
BP	0.658	0.809	0.694	0.516	0.597 (7)
BR	0.782	0.821	0.719	0.598	0.674 (3)
CT	0.690	0.371	0.522	0.391	0.485 (15)
DG	0.962	1.000	0.566	0.407	0.596 (8)
HN	0.333	0.395	0.539	0.505	0.466 (16)
KL	0.808	0.611	0.529	0.516	0.592 (9)
KO	0.386	0.379	0.638	0.865	0.686 (19)
MK	0.615	0.928	1.000	0.432	0.605 (5)
NK	0.647	0.597	0.542	0.598	0.599 (6)
PG	0.673	0.713	0.696	0.493	0.584 (10)
PL	0.454	0.862	0.333	0.404	0.423 (18)
PV	0.548	0.662	0.747	0.365	0.494 (14)
PŽ	0.858	0.860	0.472	0.370	0.527 (12)
RO	0.539	0.772	0.417	0.333	0.418 (19)
ŠA	1.000	0.822	0.762	0.454	0.659 (4)
TV	0.584	0.333	0.510	1.000	0.778 (1)
UL	0.484	0.884	0.439	0.415	0.457 (17)
ŽB	0.334	0.656	0.499	0.345	0.386 (20)

TABLE 9: Final crisp behavior score.

	Crisp				Normalized			
	Speeding	Alcohol	Telephone	Seatbelt	Speeding	Alcohol	Telephone	Seatbelt
AN	4.94	5.54	4.16	3.82	0.938	0.939	0.949	0.703
BA	5.05	5.51	4.21	3.90	0.959	0.934	0.960	0.716
BD	4.15	5.44	3.52	4.05	0.788	0.922	0.802	0.746
BP	4.94	5.82	4.12	4.57	0.939	0.987	0.939	0.841
BR	5.14	5.82	4.21	4.71	0.977	0.987	0.960	0.866
CT	4.96	5.32	3.80	3.72	0.942	0.902	0.866	0.685
DG	5.24	5.90	4.17	4.11	0.994	1.000	0.950	0.756
HN	3.94	5.34	3.79	4.80	0.748	0.905	0.864	0.884
KL	5.10	5.67	3.86	4.85	0.969	0.962	0.879	0.892
KO	4.22	5.31	4.08	5.44	0.801	0.900	0.931	1.000
MK	4.85	5.87	4.30	4.10	0.921	0.995	0.981	0.754
NK	4.83	5.66	3.98	4.69	0.918	0.960	0.907	0.862
PG	4.82	5.75	4.12	4.33	0.916	0.976	0.940	0.796
PL	4.52	5.84	3.26	4.47	0.858	0.990	0.743	0.822
PV	4.73	5.71	4.39	4.08	0.898	0.969	1.000	0.750
PŽ	5.21	5.84	3.77	3.80	0.990	0.991	0.859	0.700
RO	4.72	5.79	3.60	3.41	0.897	0.982	0.821	0.627
ŠA	5.26	5.83	4.27	4.22	1.000	0.988	0.974	0.777
TV	4.69	5.29	3.82	5.33	0.890	0.896	0.871	0.980
UL	4.57	5.85	3.59	4.60	0.869	0.992	0.819	0.845
ŽB	4.10	5.70	3.80	4.28	0.778	0.967	0.866	0.786

occur in the first half of the scale (among the better-ranked municipalities) and that the maximum shift is +7/-5 positions.

The demonstrated robustness of the novel hybrid methodology FEGRA (Fuzzy number-Entropy-GRA) confirms that it is able to create a reliable composite safety

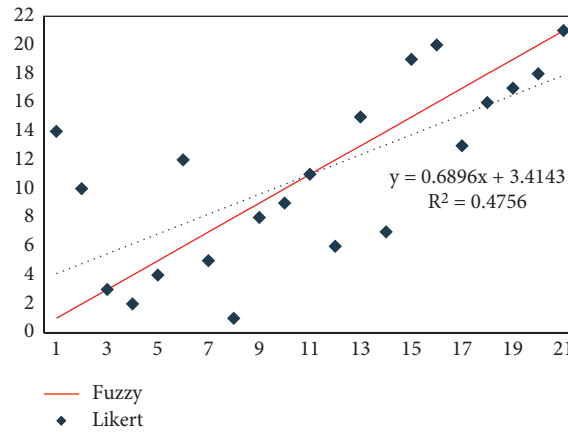


FIGURE 3: FEGRA ranking: comparison of alternative data modelling method.

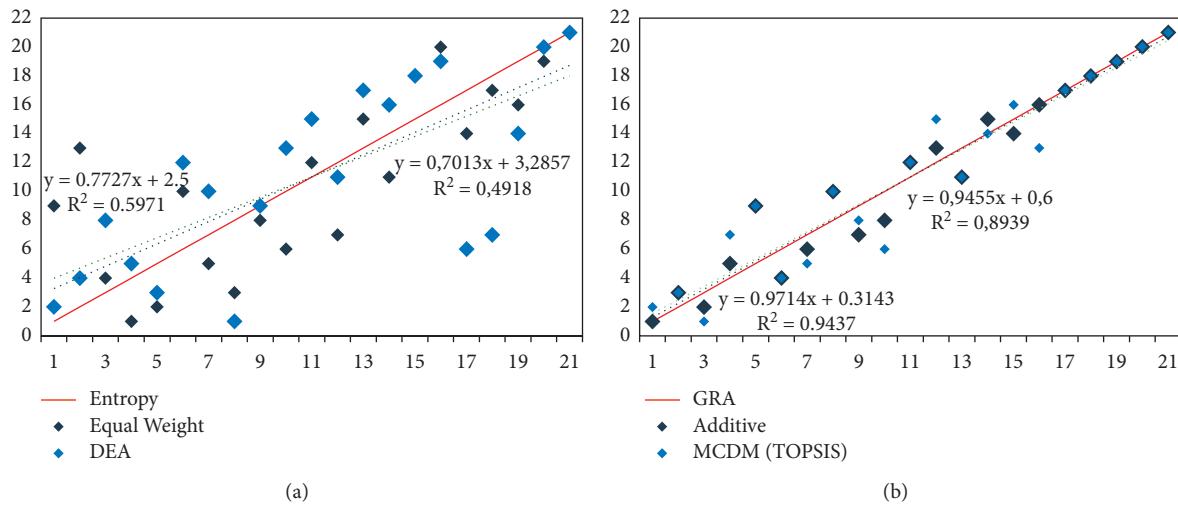


FIGURE 4: FEGRA ranking: comparison with equal and DEA weighting, (a) and comparison with additive and noncompensatory MCDM aggregation (b).

TABLE 10: Ranking shift.

	FEGRA	Data	Weighting		Aggregation	
		Likert	EW	DEA	Additive	MCDM
TV	1	13	8	1	0	1
KO	2	8	11	2	1	1
BR	3	0	1	5	-1	-2
ŠA	4	-2	-3	1	1	3
MK	5	-1	-3	-2	4	4
NK	6	6	4	6	-2	-2
BP	7	-2	-2	3	-1	-2
DG	8	-7	-5	-7	2	2
KL	9	-1	-1	0	-2	-1
PG	10	-1	-4	3	-2	-4
BA	11	0	1	4	1	1
PŽ	12	-6	-5	-1	1	3
AN	13	2	2	4	-2	-2
PV	14	-7	-3	2	1	0
CT	15	4	3	3	-1	1
HN	16	4	4	3	0	-3

TABLE 10: Continued.

	FEGRA	Data Likert	Weighting		Aggregation	
			EW	DEA	Additive	MCDM
UL	17	-4	-3	-11	0	0
PL	18	-2	-1	-11	0	0
RO	19	-2	-3	-5	0	0
ŽB	20	-2	-1	0	0	0
BD	21	0	0	0	0	0
Average rank shift		3.524	3.238	3.524	1.048	1.524
Pearson correlation		0.6896	0.7727	0.7013	0.9714	0.9454
Sig. (2-tailed)		4.15	5.31	4.29	17.83	12.64
N		0.001	0.000	0.000	0.000	0.000

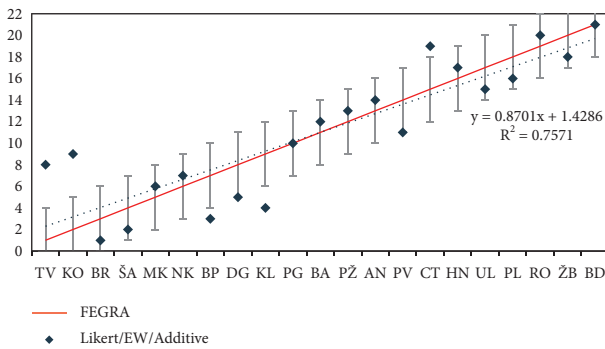


FIGURE 5: Final ranking comparison.

performance index of the territory and additionally establish a more rational ranking among the territories under evaluation.

6. Conclusions

Evaluating the state of road safety is a complicated task, especially complex in an uncertain environment. The source of uncertainty can be varied: from the selected indicators and the collection of their data, through the weighting process, to the chosen aggregation approach. This paper has just focused on assessing road safety using a composite performance index as a complex decision-making problem in uncertain circumstances. To overcome the drawbacks of data uncertainty, a systematic hybrid method—FEGRA—has been proposed. The method integrates different techniques for dealing with uncertainty at each step, namely trapezoidal fuzzy numbers for modeling indicator data, entropy-based weighting to determine the nominal importance of indicators, and grey relational analysis as a technique for aggregating weighted indicators. Data on the self-reported behavior of Montenegrin drivers collected through a questionnaire were used as indicators of safety performance (SPIs). The presented method made it possible to reliably assess traffic safety and rank municipalities in Montenegro on the basis of these behavioral SPIs.

The results show that not wearing a seatbelt is the most common negative behavior of drivers in Montenegro with the largest impact on final efficiency and ranking (i.e., it is the most important issue that policymakers should address in the future). Ranking positions are understandably very

sensitive to the “quality of presentation” of input data, especially if that data is qualitative (opinion, judgment), as here. The performed robustness analysis of FEGRA (comparative analysis) confirms that the final ranking of the territory is mostly influenced by data modeling and weighting techniques (with a total absolute average rank shift of 3.5 places), while the choice of aggregation method is, as expected, less influential (with an average rank shift of 1 rank place).

One of the advantages of the proposed FEGRA method is its efficiency and practical usefulness since it does not require complicated software, as well as the fact that it can be treated as a unique methodology for the periodical monitoring of road user behavior and the achievements of implemented measures, strategies, and policies. However, since the analysis of the proposed method showed that indicator values have a large influence on the final ranking and from the fact that the result of the fuzzy approach counts heavily on the degree that fuzzy numbers correctly represent qualitative data, the calculated results might vary based on the assigned membership function, which is the largest shortcoming of this method. Another disadvantage stems from the fact that GRA is a normalization-based technique, and the calculated results might vary based on the type of normalization. Therefore, future work should conduct a deeper sensitivity analysis of both, i.e., assigning a different membership function and applying alternative data normalization, and also explore methods that best fit data representing road safety behavior.

Data Availability

Data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] *Road Safety Strategy 2010: A Consultation Document*, National Road Safety Committee, Wellington, N.Z., 2000, <http://www.ltsa.govt.nz/publications/rs-framework.html>.
 [2] M. Koornstra, D. Lynam, G. Nilsson et al., *SUNflower: A Comparative Study of the Developments of Road Safety in*

- Sweden, the United Kingdom, and the Netherlands, SWOV, Leidschendam, 2002.
- [3] *Transport Safety Performance Indicators*, European Transport Safety Council (ETSC), Brussels, 2001.
 - [4] M. Nardo, M. Saisana, A. Saltelli, S. Tarantola, A. Hoffmann, and E. Giovannini, *Handbook on Constructing Composite Indicators: Methodology and User Guide*, OECD publishing, Paris (France), JRC47008, 2008.
 - [5] J. T. Bastos, Y. Shen, E. Hermans, T. Brijs, G. Wets, and A. C. P. Ferraz, "Traffic fatality indicators in Brazil: state diagnosis based on data envelopment analysis research," *Accident Analysis & Prevention*, vol. 81, pp. 61–73, 2015.
 - [6] A. K. Jameel and H. Evdorides, "Developing a safer road user behaviour index," *IATSS Research*, vol. 45, no. 1, pp. 70–78, 2021.
 - [7] S. Babae, M. Toloo, E. Hermans, and Y. Shen, "A new approach for index construction: the case of the road user behavior index," *Computers & Industrial Engineering*, vol. 152, Article ID 106993, 2021.
 - [8] J. Zhu and K. W. Hipel, "Multiple stages grey target decision making method with incomplete weight based on multi-granularity linguistic label," *Information Sciences*, vol. 212, pp. 15–32, 2012.
 - [9] S. de la Rosa de Saa, M. A. Gil, G. González-Rodríguez, M. T. López, and M. A. Lubiano, "Fuzzy rating scale-based questionnaires and their statistical analysis," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 1, pp. 111–126, 2015.
 - [10] A. A. Rassafi, S. S. Ganji, and H. Pourkhani, "Road safety assessment under uncertainty using a multi attribute decision analysis based on dempster-shafer theory," *KSCE Journal of Civil Engineering*, vol. 22, no. 8, pp. 3137–3152, 2018.
 - [11] X. Xiong, L. Chen, and J. Liang, "Analysis of roadway traffic accidents based on rough sets and bayesian networks," *Promet - Traffic & Transportation*, vol. 30, no. 1, pp. 71–81, 2018.
 - [12] F. Chen, J. Wang, and Y. Deng, "Road safety risk evaluation by means of improved entropy TOPSIS-RSR," *Safety Science*, vol. 79, pp. 39–54, 2015.
 - [13] M. Grdinić-Rakonjac and V. Pajković, "Evaluating the road safety of local municipalities with application of GRA: Montenegro case study," *Transactions on Transport Sciences*, vol. 11, no. 3, pp. 4–11, 2020.
 - [14] M. Castro-Nuño and M. T. Arévalo-Quijada, "Assessing urban road safety through multidimensional indexes: application of multicriteria decision making analysis to rank the Spanish provinces," *Transport Policy*, vol. 68, pp. 118–129, 2018.
 - [15] I. E. Tsoias, "Utility exchange traded fund performance evaluation. A comparative approach using grey relational analysis and data envelopment analysis modelling," *International Journal of Financial Studies*, vol. 7, no. 4, Article ID 67, 2019.
 - [16] S. M. Markabi and M. Sabbagh, "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*, vol. 7, no. 5, Article ID 25, 2014.
 - [17] S. Memiş, E. Demir, Ç. Karamaşa, and S. Korucuk, "Prioritization of road transportation risks: an application in Giresun province," *Operational Research in Engineering Sciences: Theory and Applications*, vol. 3, no. 2, pp. 111–126, 2020.
 - [18] A. Kashani, M. Bondarabadi, and M. M. Besharati, "A risk-exposure-resources approach for incorporating the performance efficiency in developing composite safety performance index," *Transportation Letters*, vol. 12, no. 7, pp. 465–470, 2020.
 - [19] H. Omrani, M. Amini, and A. Alizadeh, "An integrated group best-worst method – data envelopment analysis approach for evaluating road safety: a case of Iran," *Measurement*, vol. 152, Article ID 107330, 2020.
 - [20] B. Antić, M. Grdinić, D. Pešić, and V. Pajković, "Benchmarking of the road safety performance among the regions by using DEA," *Transportation Research Procedia*, vol. 45, pp. 78–86, 2020.
 - [21] M. Tešić, E. Hermans, K. Lipovac, and D. Pešić, "Identifying the most significant indicators of the total road safety performance index," *Accident Analysis & Prevention*, vol. 113, pp. 263–278, 2018.
 - [22] L. K. Lloyd and J. J. Forster, "Modelling trends in road accident frequency- Bayesian inference for rates with uncertain exposure," *Computational Statistics & Data Analysis*, vol. 73, pp. 189–204, 2014.
 - [23] J. Zhu, "Quantitative Models for Performance Evaluation and Benchmarking", *International Series in Operations Research & Management Science*, Springer International Publishing, Switzerland, 2014, ISBN 978-3-319-06646-2.
 - [24] P. H. Lin, "A general framework of dealing with qualitative data in DEA: a fuzzy number approach," in *Performance Measurement with Fuzzy Data Envelopment Analysis, Studies in Fuzziness and Soft Computing*, Springer-Verlag, Berlin-Heidelberg, Springer-Verlag, 2014, ISBN 978-3-642-41371-1.
 - [25] S. Tsang, B. Kao, K. Y. Yip, W.-S. Ho, and S. D. Lee, "Decision trees for uncertain data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 1, pp. 64–78, 2011.
 - [26] E. Ayati, M. A. Pirayesh Neghab, A. A. Sadeghi, and A. Mohammadzadeh Moghaddam, "Introducing roadside hazard severity indicator based on evidential reasoning approach," *Safety Science*, vol. 50, no. 7, pp. 1618–1626, 2012.
 - [27] D. Farooq, S. Moslem, R. F. Tufail et al., "Analyzing the importance of driver behavior criteria related to road safety for different driving cultures," *International Journal of Environmental Research and Public Health*, vol. 17, no. 6, Article ID 1893, 2020.
 - [28] M. Izadikhah, R. Roostae, and A. Emrouznejad, "Fuzzy data envelopment analysis with ordinal and interval data," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 29, no. 3, pp. 385–410, 2021.
 - [29] A. Barhoum and H. R. Behnood, "Composite performance indicators for helmet and seat-belt enforcement as imprecise data," *Traffic Injury Prevention*, vol. 22, no. 3, pp. 230–235, 2021.
 - [30] Y. Yang and R. John, "Grey sets and greyiness," *Information Sciences*, vol. 185, no. 1, pp. 249–264, 2012.
 - [31] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
 - [32] J.-L. Deng, "Control problems of Grey systems," *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982.
 - [33] J. Carifio and R. J. Perla, "Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes," *The Journal of Social Sciences*, vol. 3, no. 3, pp. 106–116, 2007.
 - [34] D. D. Dinh, N. H. Vu, R. C. McIlroy, K. A. Plant, and N. A. Stanton, "Examining the roles of multidimensional fatalism on traffic safety attitudes and pedestrian behaviour," *Safety Science*, vol. 124, Article ID 104587, 2020.
 - [35] B. Liu, "Why is there a need for uncertainty theory," *Journal of Uncertain Systems*, vol. 6, no. 1, pp. 3–10, 2012.

- [36] Q. Bao, D. Ruan, Y. Shen, E. Hermans, and D. Janssens, "Improved hierarchical fuzzy TOPSIS for road safety performance evaluation," *Knowledge-Based Systems*, vol. 32, pp. 84–90, 2012.
- [37] S. M. Mousavi, F. Jolai, R. Tavakkoli-Moghaddam, and B. Vahdani, "A fuzzy grey model based on the compromise ranking for multi-criteria group decision making problems in manufacturing systems," *Journal of Intelligent and Fuzzy Systems*, vol. 24, no. 4, pp. 819–827, 2013.
- [38] M.-L. Tseng and A. S. F. Chiu, "Evaluating firm's green supply chain management in linguistic preferences," *Journal of Cleaner Production*, vol. 40, pp. 22–31, 2013.
- [39] A. Awasthi and G. Kannan, "Green supplier development program selection using NGT and VIKOR under fuzzy environment," *Computers & Industrial Engineering*, vol. 91, pp. 100–108, 2016.
- [40] M. Stanković, Ž. Stević, D. K. Das, M. Subotić, and D. Pamučar, "A new fuzzy MARCOS method for road traffic risk analysis," *Mathematics*, vol. 8, no. 3, Article ID 457, 2020.
- [41] R. Rostamzadeh, A. Esmaeili, H. Sivilevičius, and H. B. K. Nobard, "A fuzzy decision-making approach for evaluation and selection of third party reverse logistics provider using fuzzy ARAS," *Transport*, vol. 35, no. 6, pp. 635–657, 2020.
- [42] A. Wersi Qazvini, H. Javanshir, and S. M. Seyedhosseni, "An application of fuzzy TOPSIS for identifying and analyzing black spots suburban areas," *Management Science Letters*, vol. 6, no. 8, pp. 561–568, 2016.
- [43] V. Pajković and M. Grdinić-Rakonjac, "Evaluation of road safety performance based on self-reported behaviour data set," *Sustainability*, vol. 13, Article ID 13837, 2021.
- [44] X. Cai, C. Lei, B. Peng, X. Tang, and Z. Gao, "Road traffic safety risk estimation method based on vehicle onboard diagnostic data," *Journal of Advanced Transportation*, vol. 2020, Article ID 3024101, 2020.
- [45] Z. Ma, C. Shao, S. Ma, and Z. Ye, "Constructing road safety performance indicators using fuzzy delphi method and grey delphi method," *Expert Systems with Applications*, vol. 38, no. 3, pp. 1509–1514, 2011.
- [46] Q.-Z. Hu, H.-P. Lu, and W. Deng, "Evaluating the urban public transit network based on the attribute recognition model," *Transport*, vol. 25, no. 3, pp. 300–306, 2010.
- [47] Y. Liu, X. Huang, J. Duan, and H. Zhang, "The assessment of traffic accident risk based on grey relational analysis and fuzzy comprehensive evaluation method," *Natural Hazards*, vol. 88, no. 3, pp. 1409–1422, 2017.
- [48] M. Grdinić-Rakonjac, B. Antić, D. Pešić, and V. Pajković, "Construction of road safety composite indicator using grey relational analysis," *Promet - Traffic & Transportation*, vol. 33, no. 1, pp. 103–116, 2021.
- [49] M. Saisana, S. Tarantola, and A. Saltelli, "Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators," *Journal of the Royal Statistical Society: Series A*, vol. 168, no. 2, pp. 307–323, 2005.
- [50] W. Becker, P. Paruolo, M. Saisana, and A. Saltelli, "Weights and importance in composite indicators: mind the gap," in *Handbook of Uncertainty Quantification*, R. Ghanem, D. Higdon, and H. Owhadi, Eds., Springer International Publishing, Switzerland, Springer International Publishing, 2017, ISBN 978-3-319-12384-4.
- [51] E. Hermans, F. Van den Bossche, and G. Wets, "Uncertainty assessment of the road safety index," *Reliability Engineering & System Safety*, vol. 94, no. 7, pp. 1220–1228, 2009.
- [52] S. Greco, A. Ishizaka, M. Tasiou, and G. Torrisi, "On the methodological framework of composite indices: a review of the issues of weighting, aggregation, and robustness," *Social Indicators Research*, vol. 141, no. 3, pp. 61–94, 2019.