

DESCRIPTION AND ROBUSTNESS OF AN AGGREGATION METHOD FOR A EUROPEAN ROAD SAFETY DATA SET

FRANK VANHOENSHOVEN, ELKE HERMANS, KOEN VANHOOF

*Hasselt University, Transportation Research Institute, Wetenschapspark 5 bus 6
3590 Diepenbeek, Belgium (elke.hermans@uhasselt.be)*

In this study, we focus on some methodological aspects linked to the performance of countries on road safety. In particular, a method aggregating different weighting methods in one, aggregated, set of weights is discussed. Moreover, the robustness of the aggregation method is tested. The proposed aggregation method, primarily working on the minimization of the differences between two rankings, proves to have a satisfactory level of robustness.

1. Introduction

1.1. Road Safety Performance in Europe

On the European level, attention is paid to the study of road safety performance. In this respect, seven essential elements are identified, i.e., alcohol and drugs, speed, protective systems, daytime running lights, vehicle, roads and trauma management (Vis, 2005; Goldenbeld & Amelink, 2006). In a second phase, appropriate safety performance indicators (SPIs) were established to enable some form of measurement for the elements stated above (Hakkert et al., 2007; Vis & Van Gent, 2007). Thirdly, a comparison between countries based on indicator data can be found in Vis & Van Gent (2007).

Two major forms of information useful for decision makers can be derived from road safety performance indicator data. First, it is interesting to know which safety performance indicator has the highest impact on road safety. The second area of concern would be a ranking of European countries in terms of road safety performance. Both the ranking of the countries and the relative importance of the safety performance indicators can be obtained by assigning appropriate weights to the seven SPIs. This will directly lead to the identification of the SPI with the highest weight, while it will enable us to construct a countries' ranking based on road safety performance. Since there are many available weighting methods, the purpose of this paper is two-fold. First, we propose a method to aggregate different weighting methods in order to get one,

aggregated, weight for each SPI. Next, we evaluate the robustness of the aggregation method.

1.2. Weighting methods

This paper takes the results of previous research as starting point. More specifically, the seven weighting methods that will be combined are depicted below. All methods use the same indicator data set consisting of seven indicators and 20 European countries. Despite the fact that there are many other weighting methods we will restrict ourselves to the following weighting methods:

- From Hermans et al. (2008): Factor Analysis (FA), Analytical Hierachy Process (AHP), Budget Allocation (BA), Data Envelopment Analysis (DEA) and Equal Weighting (EW).
- From Wang & Fan (2007): Absolute Deviation Method (ADM) and Standard Deviation Method (SDM).

The weights assigned to the seven safety performance indicators by each weighting method can be found in Table 1.

Table 1. Assigned indicator weights by different weighting methods.

Indicator	FA	AHP	BA	DEA	EW	ADM	SDM
Alcohol and drugs	0.139	0.229	0.234	0.259	0.143	0.103	0.133
Speed	0.190	0.262	0.305	0.250	0.143	0.151	0.143
Protective Systems	0.095	0.129	0.132	0.211	0.143	0.138	0.140
Daytime running lights	0.163	0.028	0.043	0.000	0.143	0.210	0.173
Vehicle	0.100	0.102	0.071	0.035	0.143	0.159	0.147
Roads	0.139	0.179	0.105	0.080	0.143	0.102	0.129
Trauma Management	0.174	0.071	0.110	0.028	0.143	0.135	0.136

2. Aggregation Methods

Various weighting methods exist. Since there does not seem to be a general consensus about which weighting method to prefer above the other ones, trying to combine the methods in an optimal way can offer a solution. In the area of data mining, the combination of different methods covering the disadvantage of one method by the advantage of another method, is typically used (Witten & Frank, 1999).

In this paper, we propose an aggregation method that both makes use of the available road safety data and an ideal representation of road safety. The latter one refers to the ranking of countries based on the number of road fatalities per

million inhabitants. This ranking can be considered to be a good expression of road safety and will consequently be used in the model. An optimization problem is formulated aimed at minimizing the differences between the ranking based on the weighted road safety indicator scores and the ideal representation. Since we approach aggregation as assigning weights to the weighting methods, aggregation can be seen as a weighting method at a higher level.

2.1. Minimizing the difference in rank

For each weighting method, a safety score per country is computed. This is easily done by multiplying the weights of each SPI by the corresponding data value and subsequently summing up these values. Once a score is obtained for each country, for each weighting method, an aggregated score for each country is computed by randomly assigning a variable weight to each weighting method. Next, the resulting countries' ranking based on these aggregated scores is constructed. For each country i , the absolute value of the difference between its position on the fatalities ranking (RC_i) and its position on the aggregated score ranking (RS_i) is calculated. The optimal set of weights is found by minimizing the sum of all those N differences. In other words, the minimization program shifts with the weights assigned to the weighting methods to minimize the total difference between the two rankings.

$$\text{Min} \sum_{i=1}^N |RS_i - RC_i|$$

2.2. Aggregated weights

The resulting aggregated weights that are assigned to each SPI based on the optimization model are shown in Table 2. It can be seen that the least weight is assigned to the daytime running lights, trauma management and vehicle indicator. The first two indicators receive most weight.

Table 2. Aggregated weights

Indicator	Weight
Alcohol and drugs	0.234
Speed	0.252
Protective Systems	0.178
Daytime running lights	0.053
Vehicle	0.083
Roads	0.120
Trauma Management	0.080

3. Robustness

An aggregation method should be robust to be of any practical use. In case the collected information about the importance of each SPI is only the result of this specific set of road safety data and not a representation of reality, we overfit the data (Witten & Frank, 1999). The robustness of the proposed aggregation method is tested by making some small adjustments to the data and subsequently examining whether those adjustments have a major impact on the outcomes. In particular, four adjustments were made before rerunning the aggregation method. The new weights will be compared to the original weights (presented in Table 2) and the effect of the changes on the ranking of countries will be discussed. The results are interpreted using different correlation coefficients, i.e., Pearson product-moment correlation coefficient, Spearman's rank correlation coefficient and Kendall's tau rank correlation coefficient.

3.1. Examining robustness

3.1.1. *Changing the ideal representation 1*

First, the impact of a change in the fatalities ranking is assessed. Therefore, the ranking of road fatalities per million inhabitants is replaced by the ranking of road fatalities per million kilometers travelled. The distance travelled by car might be a better approximation than the population size for expressing the risk of becoming fatally injured in a road accident.

3.1.2. *Changing the ideal representation 2*

Another possibility is to consider the ranking of road fatalities per 100,000 cars. A higher number of registered cars implies a higher activity on the roads. The use of this representation in the model will also be assessed.

3.1.3. *Set of countries*

Thirdly, we make some changes with respect to the countries considered in the analysis. More specifically, we focus on 15 countries and no longer consider the countries with a large discrepancy between both rankings (based on the aggregated score and on the number of road fatalities). In other words, the countries Estonia (18th vs. 12th), France (3rd vs. 8th), the Netherlands (10th vs. 3rd), Portugal (12th vs. 19th) and Slovenia (1st vs. 12th) are left out in this case.

3.1.4. *Level of detail*

Finally, the level of detail of the data is changed. Instead of using the exact indicator values, five categories are created: very low, low, medium, high and very high. Subsequently, we convert each indicator value into the corresponding 5-level value.

Table 3. Correlation coefficients expressing the robustness of the aggregation method in terms of weights

	Fatalities per mln km	Fatalities per 0.1 mln cars	Set of countries	Level of detail
Pearson	0.998	0.993	0.329	0.983
Spearman	0.995	0.983	0.321	0.986
Kendall	1.000	1.000	0.238	0.905

Table 4. Correlation coefficients expressing the robustness of the aggregation method in terms of countries' ranking

	Fatalities per mln km	Fatalities per 0.1 mln cars	Set of countries	Level of detail
Pearson	0.999	0.996	0.998	0.997
Spearman	0.995	0.982	0.986	0.985
Kendall	0.979	0.916	0.926	0.926

4. Discussion

Changing the fatalities ranking does not seem to heavily affect the weights. Table 3 indicates very high degrees of correlation between the original weights from Table 2 and the recalculated weights. The robustness of the weights implies that the countries' ranking is not largely affected by the changes neither (see Table 4). These results can be explained by the high level of harmony between the three forms of ideal representation.

The correlation coefficients with respect to the limited set of countries indicate a small degree of correlation between the sets of weights. It can be concluded that the change in the set of countries has some impact on the resulting weights. The relatively low degree of correlation in terms of weights can be justified by the fact that a particular type of countries was no longer considered in the analysis. The five countries were characterized by a notably better or worse performance in terms of road fatalities than could be expected based on their road safety score. By leaving these countries out, different results arose. Since we removed the countries causing most bias, it can be assumed that

any randomly chosen subset of countries would generate higher correlation coefficients with the original results than the correlations indicated in Table 3. At the same time, Table 4 still shows a high level of correlation with respect to the ranking of the countries.

Finally, lowering the level of detail of the data does not produce large differences. Neither the weights nor the ranking of countries seems to be affected considerably by these changes in the data set.

In general, it can be stated – based on the results in Table 3 and Table 4 – that the proposed method is a robust one. The correlation coefficients corresponding to the previous tests reveal a high level of harmony with the original results in case small adjustments are made.

5. Conclusion

In this research, an aggregation method based on differences between rankings is proposed and tested on robustness. The tests are performed on a European road safety data set.

The proposed aggregation method resulted to be a rather robust one because small changes with respect to the ideal representation (the fatalities ranking), the set of countries considered and the level of detail in the data set did not dramatically influence the resulting weights and countries' ranking.

In the future, further research can modify and sophisticate the proposed aggregation method. At the same time, other tests on robustness can be conducted and different test cases can be used.

References

1. Goldenbeld, C., & Amelink, M. (2006). Dissemination of the Road Safety Information System. In *Building the European Road Safety Observatory*. D3.4. SafetyNet
2. Hakkert, A.S., Gitelman, V., & Vis, M.A. (2006). Road Safety Performance Indicators: Theory. In *Building the European Road Safety Observatory*. D3.1. SafetyNet
3. Hermans, E., Van den Bossche, F., & Wets, G. (2008). Combining Road Safety Information in a Performance Index, AAP
4. Vis, M.A. (2005). State of the art Report on Road Safety Performance Indicators. In *Building the European Road Safety Observatory* D3.1. SafetyNet
5. Vis, M.A., & Van Gent, A. (2007) Road Safety Performance Indicators: a country comparison. In *Building the European Road Safety Observatory* D3.7a. SafetyNet
6. Wang, Y.M., & Fan, Z.P. (2007) Fuzzy preference relations: Aggregation and weight determination. *Computers & Industrial Engineering*, 53, 163-172
7. Witten, I.H., & Frank, E. (1999) DATAMINING: Practical machine learning tools and techniques with Java implementations. University of Walkatà, New Zealand