



SETTING THE WEIGHTS OF SUSTAINABILITY CRITERIA FOR THE APPRAISAL OF TRANSPORT PROJECTS

Paola Carolina Bueno Cadena, José Manuel Vassallo Magro

Transport Research Centre (TRANSYT), Technical University of Madrid, Spain

Submitted 28 November 2014; resubmitted 22 April 2015; accepted 17 June 2015

Abstract. Although the Multi-Criteria Decision Analysis (MCDA) has made progress towards appraising and measuring the performance of smart and sustainable transport projects, it still has important issues that need to be addressed such as the problem associated with incomparable quantities, the inherent subjective qualitative assessment, the complexity of identifying impacts to be included and its measurement method, and the corresponding weights. The issue of trading-off different sustainability criteria is the main unresolved matter. This problem may lead to lack of accuracy in the decision making process. This paper presents a new methodology to set the weights of the sustainability criteria used in the MCDA in order to reduce subjectivity and imprecision. We suggest eliciting criteria weights based on both expert preferences and the importance that the sustainability criteria have in the geographical and social context where the project is developed. This novel methodology is applied to a real case study to quantify sustainable practices associated with the design and construction of a new roadway in Spain. The outcome demonstrates that the approach to the weighting problem has significance and general application in a multi-criteria evaluation process.

Keywords: sustainable transport; multi-criteria analysis; decision making; criteria weighting; pairwise comparisons; REMBRANDT.

Introduction

Transport systems and infrastructure construction consume large amounts of energy and industrial products. Since the emergence of the concept of sustainability as an international priority in the 1980–1990s, there has been a growing interest in exploring strategies to promote smart and sustainable transport. Nowadays, there is a growing demand to make transport infrastructure more sustainable without compromising conventional goals (i.e., cost, quality, and schedule). Planners strive to mitigate the impacts – on the environment, the economy, and society – throughout the whole life cycle, from conception through construction, operation, maintenance, end of life processing, and final disposal.

Although there are many approaches aimed at assessing the socio-economic and environmental feasibility of transportation projects, currently there is no standardised or commonly agreed methodology offering a reliable measurement of sustainability when appraising and evaluating transport projects over their life-cycle. The available literature on smart and sustainable infrastructure (see Dasgupta, Tam 2005; Gilmour *et al.* 2011; Tsai, Chang 2012) points out that policymakers are in need of practical tools and techniques to assess sustain-

ability in all the life stages of infrastructure projects. In the EU, appraisal is generally seen as a means aimed at supporting the process of planning transport systems. The most common forms of appraisal in use in the EU member states are Cost Benefit Analysis (CBA) and Multi-Criteria Analysis (MCA) – see Bristow, Nellthorp (2000).

Although these traditionally accepted techniques offer valuable support for assessing transport projects, they hardly ever address all the components of sustainability (economic, social and environmental) in a thorough way. This research focuses on one of the most common forms of appraisal of sustainable transport systems and infrastructure: the Multi-Criteria Decision Analysis (MCDA). Despite MCDA can explicitly deal with different components of sustainability; there remain some fundamental gaps that may lead to lack of accuracy in the results of a decision making process. One of the main problems of this widely used technique for selecting the best alternative with respect to multiple criteria is precisely setting the weights to sustainability criteria in a transparent and precise way. MCDA can introduce subjectivity when evaluating the weights selected to rank different criteria.



As a new contribution to the state of knowledge, an adaptation of the traditional MCDA is developed aimed at addressing these issues and making the decision process more rational and accountable. We propose an adaptation of the weighting process that directly tackles the issues in assigning weights in a typical MCDA. To that end, we have designed a composite weighting model that allows the incorporation of consensus-based comparative judgments and preferences, along with the geographical and social context of the project.

The paper is structured as follows. Section 1 summarises the literature review on Multi-Criteria Decision Making (MCDM) for smart and sustainable transport, illustrating the knowledge base for criteria weightings. Then, the new method for obtaining preference weights is presented in section 2. Section 3 discusses the results of the application of this weighting approach to a real example of a new roadway; and last section provides a set of conclusions and final recommendations.

1. Multi-Criteria Decision Analysis for the Sustainability Appraisal of Transport Projects

Smart and sustainable transport projects are appraised in practice through a number of tools or methodological frameworks. These methods encompass the traditional analytical methodologies, as well as a number of current sustainability tools such as rating systems, frameworks, and appraisal guidelines. Among these methods, the MCA has become increasingly popular. However, as claim by Bueno *et al.* (2015), despite the fact that the current approaches offer some value for sustainability assessment, none of them can be used to carry out a holistic appraisal. Moreover, there is stillroom to improve the current tools for smart and sustainable transport infrastructure systems.

The MCDA is introduced in the following section by highlighting its strengths and weaknesses. Our aim is to examine how it works and identify whether it provides a suitable framework to integrate sustainability into existing appraisal processes.

Multi-criteria Decision Analysis

The multi-criteria technique is a suitable decision making methodology for ‘addressing complex problems featuring high uncertainty, conflicting objectives, different forms of data and information, multiple interests and perspectives, and the accounting for complex and evolving biophysical and socio-economic systems’ (Janic 2003). Its use for different purposes has been increasing over the years. There are several papers where this approach has been applied in the field of transport, for instance (Iniestra, Gutiérrez 2009; Cheng, Li 2005; Friesz *et al.* 1980; Frohwein *et al.* 1999; Giuliano 1985; Khorramshahgol, Steiner 1988). Recently, Macharis, Bernardini (2015) provided a wide overview of the increasing use of MCDA methods in the evaluation of transport projects.

Within the sustainability appraisal context, the MCDA usually includes the following main steps. First,

the alternative’s formulation and selection provided by decision makers. The second stage corresponds to the identification of sustainability criteria and the evaluation for each alternative. Sustainability criteria are defined as the basic fundamentals or principles used to judge the sustainability of transport projects and to compare the alternatives. They can be grouped into different sustainability components (economic/social/environmental). Third, the assignment of weighting coefficients to the criteria and finally, the sustainability evaluation by using a method for ranking the alternatives.

A number of authors have suggested that MCDA is the most appropriate tool to adopt decisions based on an integrated sustainability appraisal (Hyard 2012; Munda *et al.* 1998; Walker 2010). Munda (1995) says, that MCDA offers a number of advantages for policy analysis, compared with conventional economic welfare techniques. Several criteria can be taken into account simultaneously – including those difficult to monetize or quantify. Then, the technique allows capturing the full range of impacts of a project (Thomopoulos *et al.* 2009). MCDA also promotes public participation and enables stakeholder involvement.

However, despite the fact that MCDA can explicitly deal with different components of sustainability, the extensive study of multi-criteria techniques for transport projects has acknowledged issues that require further analysis including: the inherent subjective qualitative assessment, the complexity of identifying impacts to be included and its measurement method, and the obtaining of weights to criteria (Browne, Ryan 2011).

In fact, the use of weights is the main unresolved matter of this methodology. It has to do with the transparency of judgements and their influence on the final results of a multi-criteria problem. This weakness has been the subject of severe criticisms by a number of studies – see for example, Browne, Ryan (2011), Chen *et al.* (2013), Hobbs, Horn (1997), Wibowo, Deng (2011). The following section presents one of the most significant research needs that should be undertaken in order to improve the appraisal of transport projects when employing a multi-criteria approach: the use of weights and how these might be obtained in practice.

Criteria Weighting: a Gap in the Process

A number of methods for determining criteria weights in MCDA have been developed. There are thorough studies performed by different authors about weight assessment techniques – see for example by the following: Harte, Koele (1995) and Barron, Barrett (1996).

In general, there exist two weighting methods: the equal weights and the rank-order weights. The last is classified into three categories: subjective – such as pair-wise comparison, Delphi method or Analytical Hierarchy Process (AHP); objective – such as Least Mean Square (LMS), entropy method or the vertical and horizontal method; and combination weighting method – such as multiplication synthesis and additive synthesis. Wang *et al.* (2009) provided a comprehensive review of the weighting methods for MCDA.

Within this context, as Hobbs and Horn (1997) pointed out, 'the theory definitely favours trade-off judgements as a technique for choosing weights'. Macharis and Bernardini (2015) in their turn pointed out the importance of integrating decision makers in the process for obtaining weights to several criteria not yet very common in current transport projects. Some recent papers provide an overview of the theory and lessons learned from an alternative extension of the traditional MCDA where the stakeholders are explicitly taken into account during the project analysis, the Multi Actor Multi Criteria Analysis (MAMCA) (see for example, Macharis *et al.* 2012; Bergqvist *et al.* 2015; Turcksin *et al.* 2011; Sun *et al.* 2015).

In practice, the process for obtaining the relative importance of criteria might appear questionable. The 'black box' concept should be considered as an important issue since it causes a loss in credibility. In fact, 'due to a lack of procedures for aggregating the evaluations of the individual criteria and unregulated weights that were left to the whim of the decision-takers' (Sayers *et al.* 2003); some governments – such as France – have moved away from the MCDA and returned to the 'monetising approach'.

As a result, the MCA involves certain subjectiveness (Beria *et al.* 2011; Barfod *et al.* 2011). Qualitative assessment and the imputation of value-laden weightings to assumptions may lead to subjective and non-transparent biasing – see Munda (2004) and White, Lee (2009). Moreover, the fact that criteria weights are context-dependent (Ribeiro 1996) is not fully addressed in the MCDA and in practice this process may be highly questionable.

Trying to solve above-mentioned issues, some research has been conducted on the development of various approaches for criteria weighting in MCDA. Rezaei (2015), for example, proposes a new method called best-worst method that derives weights based on a pairwise comparison of the best and the worst criteria/alternatives with the other criteria/alternatives. Wang (2015) presents a fuzzy MCDM model based on a simple additive weighting method and the relative preference relation. Finally, Chen *et al.* (2014) propose an integrated weighting method that narrows the gaps between objective and subjective perspectives and offers more reasonable results.

However, despite these advances, there remains significant room for improving the setting of the weights in a practical and precise way. Novel approaches usually require complex mathematical tools, are not easy to manage or suffer from problems in modelling the subjectiveness of human decision processes. Policy makers are still in need of standardized and practical methods for evaluating the trade-offs among economic, environmental and social aspects in transport projects.

Overall, the main finding of this review is that despite the well-known strengths of the MCDA approach, it still can be improved for measuring the performance of smart and sustainable transport projects. The following section discusses a flexible approach to overcome the obstacles pointed above.

2. A Flexible Approach to Obtain Criteria Weights Incorporating the Context of the Project

The objective of this section is to present a novel process to setting the weights of sustainability criteria so as to tackle the issues previously mentioned. The novelty of the method is the separate consideration of expert preferences and the objective characteristics of the criteria in the geographical and social context of the project. The benefits of this separation are:

- the higher expected efficiency of the weighting process. In our methodology, experts were asked to state preferences among different criteria irrespective of the context and the magnitude of the impacts; then, criteria weights for sustainability items can be used for many projects because the methodology is flexible enough to adjust the weights;
- the higher expected rigorous mechanism for comparing all trade-offs among economic, environmental and social aspects. Since experts were asked to express graded comparative judgements between different criteria without having information about the project and the context; their valuation of trade-offs implies a clear representation of the extent to which the worsening of one criteria might be offset by the improving of another one;
- the higher expected objectivity of the weighting process. Our methodology considers that sustainability criteria contributes to sustainability to a greater or lesser extent depending on how sensitive they are in the context where the project is developed.

The composite model is aimed at obtaining improved weighting coefficients (designated as Improved Weights – IWs) to the sustainability criteria – see the Eq. Weights come from the context (severity level) and the consensus-based comparative judgments and preferences (Convergent Weights – CWs). These terms will be explained in greater detail below.

$$IW_i = CW_i \cdot SL_i$$

where: IW_i – Improved Weight for criterion i ; CW_i – Convergent Weight for criterion i ; SL_i – Severity Level of criterion i .

Identifying the Severity Level

As stated before, the main purpose of this step is to establish sensitivity aspects related to the geographical and social context where the project is located. In our methodology, Severity Levels (SL) for each item are obtained by adding scores achieved from the evaluation of the Present Situation (PS) and the trend in the project context. The higher the SL, the more sensitive the criteria in the context. As shown in the Eq., SL are integrated into the appraisal process by considering them part of the weighting method.

The integration of the context should be independent of the criteria magnitudes and even of the project

itself, and involves identifying the relative importance of each criterion to the sustainability appraisal in a certain context. The PS of each criterion must be evaluated using a defined scale. The analyst then compares the value for the context with an acknowledged reasonable value for this specific criterion (defined as the average value for other similar contexts). It allows policy makers to put the environmental, economic and social performance of the region where the project will be implemented into context, by ‘benchmarking’ them to other countries with similar geographic, social or regional characteristics.

The criterion is likely to have a greater impact on global sustainability if the PS is considered to be much worse than the acknowledged reasonable value of similar social and economic context. Therefore, a score is allocated to the PS for each criterion in context according to Fig. 1. In this case, the question of what is much/ slightly/moderately better or worse could be answered by the current state of knowledge or legislation in the particular discipline of the criteria.

A short example may serve to illustrate the process of evaluating of the PS presented in previous paragraphs. Imagine two transport projects with the same characteristics but implemented in different countries: Germany and Spain. To solve the problem of understanding the importance of the social sustainability item ‘employment effects’ (i.e. the PS for this criterion) for both projects, we need to compare the unemployment rate in Germany and Spain with the average value for different European countries. According to the World Bank database, the percentage of the labour force that is without work – and is actively looking for work – is 5.3% for Germany and 26% for Spain, whereas the average of the unemployment rate in European countries is 10.9%. In the case of the project to be developed in Germany, we can reasonably assume that the PS for this criterion is much better than the context average, and 0 points should be assigned. In contrast, in the case of Spain, the PS is much worse and, according to Fig. 1, 5 points should be given to this criterion.

To calculate the level of severity our methodology proposes to evaluate also the trend, in addition to the PS, for each criterion in the geographical context where the project is located. The main outcome of the present task is the classification of each item trend as ‘improving’, ‘stable’ or ‘worsening’ and the allocation of the corresponding score (0 points, 1 point and 2 points, respectively).

Continuing with the example described above, the percentage of the unemployment rate in Spain was increasing at the time of appraising the project while in Germany was decreasing for many decades. Then, for the first case, the criterion trend should be classified as ‘worsening’ and a total of 2 points should be allocated. In contrast, for the case of Germany, 0 points should be assigned. In summary, the ‘employment effects’ item is more sensitive and it has higher level of importance for a project to be implemented in Spain and then, according to the Eq., it should have a higher weight than the same item for the German project.

In order to conduct a thorough evaluation of the project, the previously described process for the ‘employment effects’ item should be repeated for each one of the set of major sustainability items to be considered for the project over its life-cycle.

Obtaining Convergent Weights

As recognised by Gühnemann *et al.* (2012), the weight allocated to each criterion and sub-criterion in the framework should also reflect the decision-makers’ preferences. Since it aims to narrow the gap between theoretical sustainable requirements, current design practices and decision-making processes, it is considered crucial to incorporate the decision-makers’ preferences irrespective of the context. In order to incorporate decision-makers preferences, we propose to use a combination of the Ratio Estimation in Magnitudes or decibels to Rate Alternatives which are Non-Dominated system (REMBRANDT system); and the Delphi method.

A pairwise comparison method is required to determine the weights for each criterion in order to establish a trade-off between different criteria. We use the REMBRANDT technique to derive weights, since it is a further development of the well-known original *Analytic Hierarchy Process (AHP)*. As consensus is rarely reached in practice, the Delphi technique – see Linstone, Turoff (1975) – should be used to achieve a convergence of opinion from experts. The process can be completed throughout the following stages:

- (1) *Questionnaire design.* Pairwise comparisons are organised based on a previously identified criteria list. Experts are then asked to compare the importance of different sustainability criteria based on a –8 to +8 scale known as the REMBRANDT scale – see Olson *et al.* (1995).

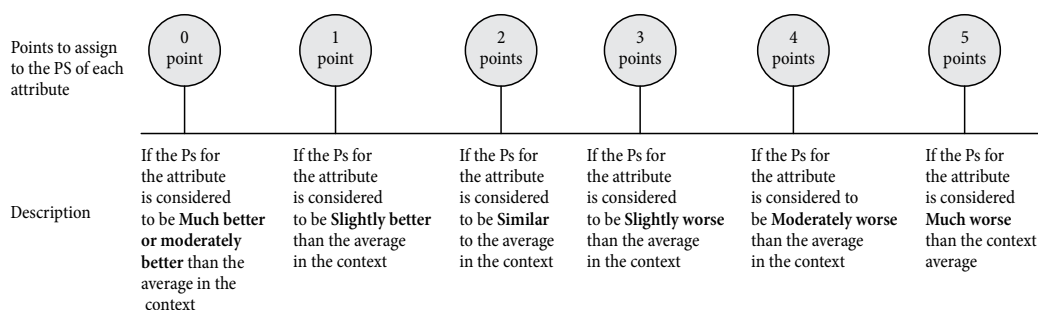


Fig. 1. Evaluation of the Present Situation (PS) for each attribute

- (2) *Conducting a survey.* A number of experts are selected to complete the questionnaire. The survey needs to reflect the views of as many interested parties as practicable. A minimum of 30 respondents is required for this weighting exercise to be robust.
- (3) *REMBRANDT calculations.* Each expert surveyed has to complete a matrix of preferences. Each element of the matrix represents the preferences stated by the expert. Criteria weights are then obtained using the REMBRANDT technique – see Olson *et al.* (1995).
- (4) *Statistical test.* A statistical test is conducted to evaluate the convergence of opinion for a weighting process to be deemed robust and valid. For this methodology, we developed a simulation based on a cross validation technique to estimate the level of consensus among the panel of experts. For this purpose, we used the R software (<http://www.r-project.org>).

The test consisted in dividing the data of the experts surveyed (weights obtained from experts) into two equal-sized parts. The test then compared the answers of both groups in order to find significant differences. This procedure was repeated 1000, 10000 and 100000 times with randomly selected groups. The result was a p-value distribution for each criterion. The p-value was used to analyse the data set and test the null hypothesis (Ho): both groups' answers are significantly different. We adopted a 5% significance, thus accepting Ho if the p-value was 0.05 or lower. If it were higher than 5%, we would not have enough evidence to assume there are significant different answers, and would therefore reject Ho.

If the level of consensus is sufficient (p-values higher than 0.05), a Delphi method will not be necessary. In this case, the average of the weights obtained from the survey is used as the Convergent Weights (CWs). Otherwise, if the statistical test is unable to prove the required level of consensus, the Delphi technique is applied to achieve a convergence of opinion on weight estimations until CWs are obtained. Multiple interactions are used to achieve consensus for the panel of experts. The procedure for this second round is summarised in the following steps:

- *Step 1.* A summary of the general result is returned to the experts surveyed, allowing them to revise their judgements or specify their reasons for remaining outside the consensus.
- *Step 2.* The procedure specified in (3) and (4) is repeated; that is, the REMBRANDT calculations for criteria weights and the statistical test.
- *Step 3.* The iterative process can be stopped once consensus is achieved.

3. An Example Applied to a Transport Project

To demonstrate the feasibility and usefulness of the proposed methodology, this section describes its application to a decision-making case study concerning the design and construction of a new interurban roadway in Spain.

First, appropriate criteria to measure the performance of the different alternatives were identified – taking into account economic efficiency, environmental protection, and social aspects. The list of criteria for this real case study is shown in the Table, second column.

Beyond previously identified sustainability criteria, we had to identify the attribute in context to evaluate the performance for each criterion. For example, in order to have information about the importance for the 'investment costs' criterion, we may evaluate the 'budgetary availability for infrastructure spending' attribute in context. The Table (third column) also presents the list of criteria to be evaluated in context.

To help better illustrate how to incorporate the context into the process, the Table presents a complete evaluation of the PS and the trend for each criterion in the particular geographical area where the project is going to be located (Spain). This information was obtained mainly from the World Bank database, the Eurostat data source yearbook and other official sources. We allocated a score to the PS for each criterion in its context (*shown in italics*, below the 'average in context' value). The trends for each particular criterion were classified, and a score was allocated to each attribute trend (*shown in italics*). Finally, the severity level was obtained by adding together the scores for the PS and the scores from the trend.

Taking the example of 'employment effects' explained above, we obtained a severity level of 7 points to this item – see the Table. This result comes up of adding the points assigned to the PS of this criterion in the context of Spain (5 points) plus the points given for revealing a worsening trend (2 points).

Another example can be given for the 'distributive effects of the project' criterion. We took advantage of the information provided by Eurostat for the most widely used measure of income inequality, the Gini coefficient. The value of this coefficient in Spain was around 0.34, while the average of the EU countries was 0.29. Since the PS is considered to be moderately worse than the average value in context, a score of 4 points will be ultimately assigned to this item according to Fig. 1 (see the Table). In addition, since the trend for the Gini coefficient is stable, 1 point will be assigned to this criterion. As a result, a severity level of 5 points was obtained for the 'distributive effects of the project' criterion.

This means that, at the time of conducting this analysis, the 'distributive effects of the project' item is likely to have less importance in the context of Spain than the 'employment effects'. As a consequence, assuming the same CW for both criteria and according to the Eq., the 'employment effects' item should have a higher final weight (IW).

To obtain the preference weights, we organised pairwise comparisons based on the list of previously identified criteria in order to consider the opinions of stakeholders for a sound decision process. We asked 250 experts to complete a questionnaire to determine priorities among the different criteria related to roadways throughout their life cycle. The survey included experts from transportation research centres, public sector

Table. Present Situation (PS) and trends for attributes in context

Sustainability component	Criterion	Criterion to evaluate in context	Units	Present situation	Average in context	Trend***	Severity level (0–7 POINTS)	Improved weights (IW)
Economic	Investment costs	Budgetary availability for infrastructure spending	%	2	5 (4)	↘ (2)	6	5
	Maintenance costs							5
	Road operating costs							5
	Vehicle operating costs							4
	Accident costs	Road accident rates	Deaths per 100,000 inhabitants	5.4	4 (4)	↘ (2)	6	17
Travel time	Average congestion level	16	23 (1)	– (1)	2	2		
Environmental	Energy consumption*	Energy consumption	GJ per year	135.59	163.86 (2)	↗ (0)	2	2
	Fuel consumption	Energy consumption	Kilograms of oil equivalent per person per year	3228	3985 (2)	– (1)	3	3
	CO ₂ emissions**	CO ₂ emissions	Metric tons per capita	5.85	7.32 (1)	↗ (0)	1	1
	Habitat fragmentation and negative effects on species	Environmental fragility of the habitat	(1–5)	5	3 (5)	– (1)	6	6
	Landscape degradation/negative visual impacts							7
	Noise pollution	Noise levels	Population exposed to traffic noise above 65 dB	27.7%	30% (2)	↘ (2)	4	4
Community disruption	Aggregated indicator of community cohesion	35%	70% (5)	– (1)	6	10		
Social	Impacts on businesses and community services	Fragility of business environment	(1–5)	4	2 (4)	↘ (2)	6	11
	Employment effects	Unemployment rate	%	–26.3%	–11.2% (5)	↘ (2)	7	11
	Distributive effects of the project	Gini coefficient	Dimensionless quantity	0.34	0.29 (4)	– (1)	5	8

Notes: *Including total energy consumed in the construction-maintenance (i.e. extraction of materials and resources), and operating phase (i.e. fossil fuel energy consumption); **Including the carbon embodied in construction materials, fossil fuels and construction machinery vehicles (construction and maintenance) and the carbon embodied in vehicles and fossil fuels and direct emissions due to combustion of fossil fuels (operation); ***According to the methodology, classified as: improving (↗), stable (–) or worsening (↘).

managers, specialists from international organisations (the World Bank and the European Investment Bank, among others), as well as professors, researchers, designers and practitioners. In the questionnaire, respondents were asked to state their preferences for each pairwise comparison and mark their choices, presented as the REMBRANDT scale. The questionnaire explicitly asked to answer about the criteria irrespective of the context of the project and the magnitude of the impacts. The specific project to be evaluated was not mentioned in the questionnaire.

Based on their personal view of the relative importance of the economic, environmental and social criteria defined for this case study, we obtained criteria weights by using the REMBRANDT technique – see Fig. 2. The decision-makers showed a strong preference

for accident cost savings over other criteria. Differences were not very significant for environmental and social criteria. A weak preference was found for all social criteria over environmental and even economic criteria in terms of sustainability. This implies that the worsening of a sustainability item could be offset by improving a social aspect. Nevertheless, the magnitude of the impact is important in the final sustainability evaluation.

Finally, the level of consensus among the panel of experts was estimated by using the statistical test based on a cross validation technique developed by our methodology. We divided the data into two parts and compared the answers of the experts surveyed in both groups. We repeated the procedure 1000, 10000 and 100000 times, and finally obtained a p-value distribution for each sustainability criterion – see Fig. 3.

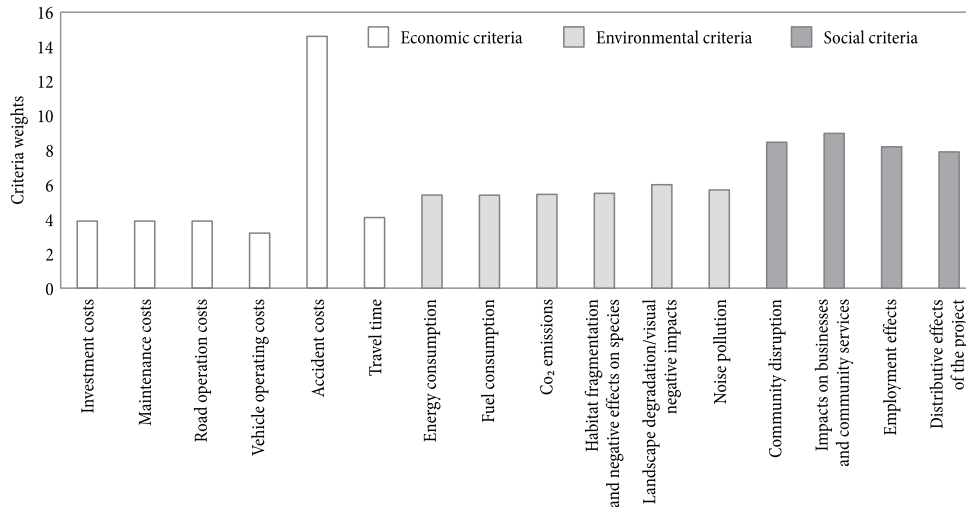


Fig. 2. Summary of results from the weighting survey

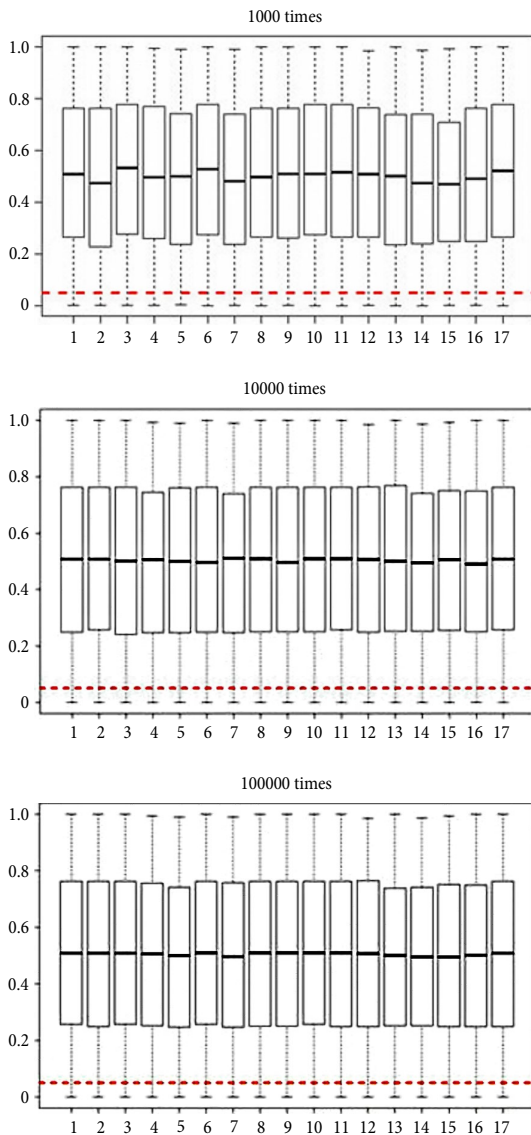


Fig. 3. P-value distributions

These results showed reported p-values higher than 0.05, allowing us to conclude that the level of consensus is good enough so the implementation of a Delphi method was not necessary at the end.

As indicated in the statistical process, we assumed criteria weights derived from the weighting survey as the final CWs. These were subsequently adjusted with the severity level. Finally, an IW was found for each sustainability criterion by applying the Eq. Results are shown in the last column of the Table.

Conclusions and Discussion of Findings

In this paper, we proposed a new and transparent method for effectively assisting decision makers in determining the criteria weightings for transport project appraisal.

This approach obtains the weights by considering separately the sensitivity of the criteria in the geographical context where the project is situated, and the trade-offs among different criteria from consensus-based comparative judgments and preferences. To show the applicability of the methodology proposed in this paper, we applied it to a real road project in Spain. The practical implementation of this approach demonstrated that it is suitable for setting the weights of different sustainability criteria in roadway projects. This example shows that the proposed weighting method has a number of advantages including:

- its simplicity and comprehensibility; this facilitates the understanding and usage of this method for practical applications;
- its flexibility and ability of replication; it can be adapted to different real-world applications. It also has the advantage of allowing, for many road projects, the use of the same consensus-based comparative judgments and preferences results obtained from the survey conducted in this research;

- its ability to increase the efficiency, rigor and objectivity of the process aimed at setting the weights; it adequately handles the subjectivity of the process of trading-off different sustainability criteria.

Future research should continue testing the applicability and usefulness of the proposed methodology by applying it to other real-word case studies and comparing the results with other approaches for setting weightings. We also suggest improving the method's validation by applying it to projects that were already appraised to compare the results of our methodology ex-post. This way, some relevant questions such as what would have happened if we had proceeded according to other different weight method (for example, with the final selection of the best alternative) can be answered.

Finally, the novel weighting method is expected to have significance and great potential to be implemented in the multi-criteria evaluation process aimed at assessing sustainability. However, much research should be conducted in order to solve other well-known issues of the MCDA such as the inter-temporal aggregation of environmental, social and economic impacts to improve the life-cycle evaluation.

References

- Barfod, M. B.; Salling, K. B.; Leleur, S. 2011. Composite decision support by combining cost-benefit and multi-criteria decision analysis, *Decision Support Systems* 51(1): 167–175. <http://dx.doi.org/10.1016/j.dss.2010.12.005>
- Barron, F. H.; Barrett, B. E. 1996. Decision quality using ranked attribute weights, *Management Science* 42(11): 1515–1523. <http://dx.doi.org/10.1287/mnsc.42.11.1515>
- Bergqvist, R.; Macharis, C.; Meers, D.; Woxenius, J. 2015. Making hinterland transport more sustainable a multi actor multi criteria analysis, *Research in Transportation Business & Management* 14: 80–89. <http://dx.doi.org/10.1016/j.rtbm.2014.10.009>
- Beria, P.; Maltese, I.; Mariotti, I. 2011. Comparing cost benefit and multi-criteria analysis : the evaluation of neighbourhoods' sustainable mobility, in *XIII Riunione Scientifica della Società Italiana degli Economisti dei Trasporti e della Logistica (SIET)*, 2011, Messina, Italy, 1–26.
- Bristow, A. L.; Nellthorp, J. 2000. Transport project appraisal in the European Union, *Transport Policy* 7(1): 51–60. [http://dx.doi.org/10.1016/S0967-070X\(00\)00010-X](http://dx.doi.org/10.1016/S0967-070X(00)00010-X)
- Browne, D.; Ryan, L. 2011. Comparative analysis of evaluation techniques for transport policies, *Environmental Impact Assessment Review* 31(3): 226–233. <http://dx.doi.org/10.1016/j.eiar.2010.11.001>
- Bueno, P. C.; Vassallo, J. M.; Cheung, K. 2015. Sustainability assessment of transport infrastructure projects: a review of existing tools and methods, *Transport Reviews* 35(5): 622–649. <http://dx.doi.org/10.1080/01441647.2015.1041435>
- Chen, S.; Leng, Y.; Mao, B.; Liu, S. 2014. Integrated weight-based multi-criteria evaluation on transfer in large transport terminals: a case study of the Beijing south railway station, *Transportation Research Part A: Policy and Practice* 66: 13–26. <http://dx.doi.org/10.1016/j.tra.2014.04.015>
- Chen, Y.; Yu, J.; Khan, S. 2013. The spatial framework for weight sensitivity analysis in AHP-based multi-criteria decision making, *Environmental Modelling & Software* 48: 129–140. <http://dx.doi.org/10.1016/j.envsoft.2013.06.010>
- Cheng, E. W. L.; Li, H. 2005. Analytic network process applied to project selection, *Journal of Construction Engineering and Management* 131(4): 459–466. [http://dx.doi.org/10.1061/\(ASCE\)0733-9364\(2005\)131:4\(459\)](http://dx.doi.org/10.1061/(ASCE)0733-9364(2005)131:4(459))
- Dasgupta, S.; Tam, E. K. L. 2005. Indicators and framework for assessing sustainable infrastructure, *Canadian Journal of Civil Engineering* 32(1): 30–44. <http://dx.doi.org/10.1139/l04-101>
- Friesz, T. L.; Tourreilles, F. A.; Han, A. F.-W.; Fernandez, J. E. 1980. Comparison of multicriteria optimization methods in transport project evaluation (abridgement), *Transportation Research Record: Journal of the Transportation Research Board* 751: 38–41.
- Frohwein, H. I.; Lambert, J. H.; Haimes, Y. Y.; Schiff, L. A. 1999. Multicriteria framework to aid comparison of roadway improvement projects, *Journal of Transportation Engineering* 125(3): 224–230. [http://dx.doi.org/10.1061/\(ASCE\)0733-947X\(1999\)125:3\(224\)](http://dx.doi.org/10.1061/(ASCE)0733-947X(1999)125:3(224))
- Gilmour, D.; Blackwood, D.; Banks, L.; Wilson, F. 2011. Sustainable development indicators for major infrastructure projects, *Proceedings of the Institution of Civil Engineers – Municipal Engineer* 164(1): 15–24. <http://dx.doi.org/10.1680/muen.800020>
- Giuliano, G. 1985. A multicriteria method for transportation investment planning, *Transportation Research Part A: General* 19(1): 29–41. [http://dx.doi.org/10.1016/0191-2607\(85\)90004-4](http://dx.doi.org/10.1016/0191-2607(85)90004-4)
- Gühnemann, A.; Laird, J. J.; Pearman, A. D. 2012. Combining cost-benefit and multi-criteria analysis to prioritise a national road infrastructure programme, *Transport Policy* 23: 15–24. <http://dx.doi.org/10.1016/j.tranpol.2012.05.005>
- Harte, J. M.; Koele, P. 1995. A comparison of different methods for the elicitation of attribute weights: structural modeling, process tracing, and self-reports, *Organizational Behavior and Human Decision Processes* 64(1): 49–64. <http://dx.doi.org/10.1006/obhd.1995.1089>
- Hobbs, B. F.; Horn, G. T. F. 1997. Building public confidence in energy planning: a multimethod MCDM approach to demand-side planning at BC gas, *Energy Policy* 25(3): 357–375. [http://dx.doi.org/10.1016/S0301-4215\(97\)00025-6](http://dx.doi.org/10.1016/S0301-4215(97)00025-6)
- Hyard, A. 2012. Cost-benefit analysis according to Sen: An application in the evaluation of transport infrastructures in France, *Transportation Research Part A: Policy and Practice* 46(4): 707–719. <http://dx.doi.org/10.1016/j.tra.2012.01.002>
- Iniestra, J. G.; Gutiérrez, J. G. 2009. Multicriteria decisions on interdependent infrastructure transportation projects using an evolutionary-based framework, *Applied Soft Computing* 9(2): 512–526. <http://dx.doi.org/10.1016/j.asoc.2008.07.006>
- Janic, M., 2003. Multicriteria evaluation of high-speed rail, Transrapid Maglev and air passenger transport in Europe, *Transportation Planning and Technology* 26(6): 491–512. <http://dx.doi.org/10.1080/0308106032000167373>
- Khorramshahgol, R.; Steiner, H. M. 1988. Resource analysis in project evaluation: a multicriteria approach, *Journal of the Operational Research Society* 39(9): 795–803. <http://dx.doi.org/10.1057/jors.1988.139>

- Linstone, H. A.; Turoff, M. 1975. *Delphi Method: Techniques and Applications*. Addison-Wesley Educational Publishers Inc. 621 p.
- Macharis, C.; Bernardini, A. 2015. Reviewing the use of multi-criteria decision analysis for the evaluation of transport projects: time for a multi-actor approach, *Transport Policy* 37: 177–186. <http://dx.doi.org/10.1016/j.tranpol.2014.11.002>
- Macharis, C.; Turcksin, L.; Lebeau, K. 2012. Multi actor multi criteria analysis (MAMCA) as a tool to support sustainable decisions: state of use, *Decision Support Systems* 54(1): 610–620. <http://dx.doi.org/10.1016/j.dss.2012.08.008>
- Munda, G. 2004. Social multi-criteria evaluation: methodological foundations and operational consequences, *European Journal of Operational Research* 158(3): 662–677. [http://dx.doi.org/10.1016/S0377-2217\(03\)00369-2](http://dx.doi.org/10.1016/S0377-2217(03)00369-2)
- Munda, G. 1995. *Multicriteria Evaluation in a Fuzzy Environment: Theory and Applications in Ecological Economics*. Physica-Verlag HD. 257 p. <http://dx.doi.org/10.1007/978-3-642-49997-5>
- Munda, G.; Nijkamp, P.; Rietveld, P. 1998. Environmental decision making: a comparison between cost-benefit analysis and multicriteria decision aid, in S. Faucheux, M. O'Connor, J. Van der Straaten (Eds). *Sustainable Development: Concepts, Rationalities and Strategies*, 227–238. http://dx.doi.org/10.1007/978-94-017-3188-1_14
- Olson, D. L.; Fliedner, G.; Currie, K. 1995. Comparison of the REMBRANDT system with analytic hierarchy process, *European Journal of Operational Research* 82(3): 522–539. [http://dx.doi.org/10.1016/0377-2217\(93\)E0340-4](http://dx.doi.org/10.1016/0377-2217(93)E0340-4)
- Rezaei, J. 2015. Best-worst multi-criteria decision-making method, *Omega: the International Journal of Management Science* 53: 49–57. <http://dx.doi.org/10.1016/j.omega.2014.11.009>
- Ribeiro, R. A. 1996. Fuzzy multiple attribute decision making: a review and new preference elicitation techniques, *Fuzzy Sets and Systems* 78(2): 155–181. [http://dx.doi.org/10.1016/0165-0114\(95\)00166-2](http://dx.doi.org/10.1016/0165-0114(95)00166-2)
- Sayers, T. M.; Jessop, A. T.; Hills, P. J. 2003. Multi-criteria evaluation of transport options – flexible, transparent and user-friendly?, *Transport Policy* 10(2): 95–105. [http://dx.doi.org/10.1016/S0967-070X\(02\)00049-5](http://dx.doi.org/10.1016/S0967-070X(02)00049-5)
- Sun, H.; Zhang, Y.; Wang, Y.; Li, L.; Sheng, Y. 2015. A social stakeholder support assessment of low-carbon transport policy based on multi-actor multi-criteria analysis: the case of Tianjin, *Transport Policy* 41: 103–116. <http://dx.doi.org/10.1016/j.tranpol.2015.01.006>
- Thomopoulos, N.; Grant-Muller, S.; Tight, M. R. 2009. Incorporating equity considerations in transport infrastructure evaluation: current practice and a proposed methodology, *Evaluation and Program Planning* 32(4): 351–359. <http://dx.doi.org/10.1016/j.evalprogplan.2009.06.013>
- Tsai, C. Y.; Chang, A. S. 2012. Framework for developing construction sustainability items: the example of highway design, *Journal of Cleaner Production* 20(1): 127–136. <http://dx.doi.org/10.1016/j.jclepro.2011.08.009>
- Turcksin, L.; Macharis, C.; Lebeau, K.; Boureima, F.; Van Mierlo, J.; Bram, S.; De Ruyck, J.; Mertens, L.; Jossart, J.-M.; Gorissen, L.; Pelkmans, L. 2011. A multi-actor multi-criteria framework to assess the stakeholder support for different biofuel options: the case of Belgium, *Energy Policy* 39(1): 200–214. <http://dx.doi.org/10.1016/j.enpol.2010.09.033>
- Walker, G. 2010. Environmental justice, impact assessment and the politics of knowledge: the implications of assessing the social distribution of environmental outcomes, *Environmental Impact Assessment Review* 30(5): 312–318. <http://dx.doi.org/10.1016/j.eiar.2010.04.005>
- Wang, J.-J.; Jing, Y.-Y.; Zhang, C.-F.; Zhao, J.-H. 2009. Review on multi-criteria decision analysis aid in sustainable energy decision-making, *Renewable and Sustainable Energy Reviews* 13(9): 2263–2278. <http://dx.doi.org/10.1016/j.rser.2009.06.021>
- Wang, Y.-J. 2015. A fuzzy multi-criteria decision-making model based on simple additive weighting method and relative preference relation, *Applied Soft Computing* 30: 412–420. <http://dx.doi.org/10.1016/j.asoc.2015.02.002>
- White, L.; Lee, G. J. 2009. Operational research and sustainable development: tackling the social dimension, *European Journal of Operational Research* 193(3): 683–692. <http://dx.doi.org/10.1016/j.ejor.2007.06.057>
- Wibowo, S.; Deng, H. 2011. Intelligent decision support for criteria weighting in multicriteria analysis for evaluating and selecting cargo ships under uncertainty, in *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS 2011)*, 16–18 March 2011, Hong Kong, China, 2: 1411–1416.