Comparisons of robustness measures as a communicative means for involvement of decision makers

A. V. Jensen
Department of Transport, Technical University of Denmark, Denmark

Abstract

Decisions about infrastructure projects or new policies in the transport sector have traditionally been based on cost benefit analysis (CBA). However, as society in general becomes more and more complex, this affects the decision-making process. Thus decision-makers are confronted with the difficult problem of evaluating potential outcomes and choosing policies to achieve the desired outcomes in the presence of this complexity. In this respect, multi-criteria decision analysis (MCDA) becomes a useful tool for the decision-makers as this type of analysis is able to perform an assessment based on a more comprehensive evaluation framework by also taking into account non-quantifiable impacts. This article concerns decision-making relating to transport projects involving multiple objectives (MCDA); especially it addresses how to measure the robustness of these decisions as regards involving views of multiple and diversified stakeholders within the MCDA. The communicative means for involving stakeholders and decision-makers in the decision process are also discussed. Specifically, based on theory and case studies, a comparison of different measures for decision robustness are treated including also how these measures can be communicated to the decision-makers. Furthermore, it is examined how the choice of MCDA methodology can affect the robustness. Finally, in addition to summarising the findings, some recommendations for applying robustness measures are given.
Keywords: sensitivity analysis, multi-criteria decision analysis, stakeholders, transport appraisal, rank order distribution, ordinal ranking.

1   Introduction

Today's decision-makers (DMs) are expected to be socio-economic responsible when taking decision – but also to be social responsible with regards to sustainability, equity and so on. In most real-world policy situations there are many alternatives, many uncertainties, many stakeholders and many consequences of interest (Walker [1]). This and the fact that there is usually no single decision-maker and getting agreement on a single set of preferences is probably not possible. In this respect multi-criteria decision analysis (MCDA) becomes a useful tool for the decision-makers. MCDA should not be seen as a prescriptive answer (often there is no optimal solution), but a transparent and informative decision process which helps to uncover how peoples' intuitive decision procedures can be informed by a structured rational analytic process (Ananda and Herath [2]). Thereby the purpose of MCDA is to assist decision-makers in choosing a course of action from among complex alternatives. The use of the term 'assist' emphasizes that MCDA is used as a decision aid and is not intended to replace the judgement of the decision-makers.

In public decision-making the inclusion of various stakeholders has become a natural part of the process. The process of public decision-making have to deal with two difficulties: the difficulty of instituting a structured public debate that can serve to legitimize the decisions made and the difficulty of providing appropriate instruments for evaluating investment projects that are transparent for all (Damart and Roy [3]). This lead to the fact that scientists cannot provide any useful input without interacting with the rest of society and the rest of the society cannot perform any sound decision making without interacting with the scientists (Munda [4]). As stated earlier in most real-world policy situations there are many uncertainties and many stakeholders, all with their own preferences. In order to better deal with these two aspects of decision-making (Funtowicz and Ravetz [5]) have developed an epistemological framework called 'post-normal science". Post-normal science can be described by Figure 1. When both
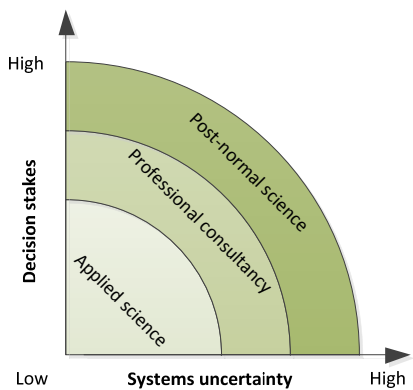

Figure 1:   Graphical representation of post-normal science.

uncertainty and stakes are small, we are in the realm of normal academic science; here it is safe to rely on standard routines and procedures. When either uncertainty or stakes are in the medium range, then the application of routine techniques and standardized knowledge is no longer enough. In these cases, skill and judgement are required to adjust the "general knowledge" available to the "Special situation" by professional consultancy. Lastly, when the stakes and systems uncertainty are very high post-normal science occur. Petersen *et al*. [6] give a description of how the post-normal science paradigm has been put into practice in the Netherlands.

1.1 Sensitivity analysis

Sensitivity analysis (SA) aims at determining how the output of a quantitative analysis depends on the inputs and thereby assessing the robustness. In MCDA, SA could be a means of explaining to the DM the implications and possible inconsistencies of his judgements. However, in decision problems with high awareness from the public, like transport policy, SA could also be a means for taking into account various different preferences.

In MCDA the weights assigned to the decision criteria represent the importance of the criteria. When criteria cannot be expressed in quantitative terms (such as cost, weight, volume, etc.), it is difficult to represent accurately the importance of these criteria. In a situation like this, the decision-making process can be improved considerably by identifying the critical criteria (those who have the potential to alter the ranking of the alternatives) and then re-evaluating more accurately the weights of these criteria. The intuitive belief is that the criterion with the highest weight is the most critical. This may not always be true and, in some particular instances, the criterion with the lowest weight may be the most critical one (Triantaphyllou and Sánchez [7]).

Several different publications have suggested various methods for assessing the robustness of decisions by applying SA. Many have found that to examine the differences in preference weights are of high importance [7–9]. They all acknowledge that variations in the criteria weights can have big influences on the results of a MCDA. Recently Jessop [10] has suggested a method for modelling uncertainties using probabilistic weight, Rios Insua and French [11] propose an SA method to find the competitors of a current optimal alternative. In Butler *et al*. [12] and Butler and Olson [13] a method for doing simulation over the criteria weights while the rank order weights on the measures is maintained, but the weights are otherwise generated at random.

Providing decision support in the public domain with high awareness from the public and active involvement of various stakeholders groups in the decision process put demands on the technique used. The techniques to be used have to be fast, transparent and clear and at the same time be valid in its methodology. This applies to decision support in transport planning, where there is an increasing tendency to involve stakeholders in the process of planning and decision support. This paper aims at describing how to assess the robustness of decision support in a transparent, fast and operational way to be used during the decision process directly with the stakeholders. So far this has had little attention in the literature,

but is important if the decision-making shall be able to deal with the complexity of today's society.

This paper is structured as follows; after this introduction a description of methods for assessing the robustness of decisions supported by MCDA is given, thereafter follows a demonstration of a specific decision problem supported by the use of MCDA and a measurement of the robustness of the decision aid. Finally, a conclusion and some perspectives are given.

2 Assessing the robustness by sensitivity analysis

French *et al.* [14] argue that sensitivity analysis (SA) can support the decision process in eight ways: exploration of consequence models, exploration of the relationship between science and the consequence model, support the elicitation of judgemental inputs, development of efficient computational algorithms, design of surveys and experiments to gather relevant data, guidance for inferences, forecasts and decisions, communication and consensus building and development of understanding.

All of the above listed reasons for SA are relevant for public decision, but especially communication and consensus building are important. In order to get both public and political support for the decisions it is essentially that the foundation for the decision-making process can be made transparent and thereby easily to communicate. Furthermore the existing of several different preferences call for a methodology to examine how the decisions are sensitive to differences in preferences.

There are several other reasons for conducting SA especially in a MCDA framework. First is because of the nature of MCDA process, which inherently contains various levels of uncertainty because of qualitative and subjective choices of different parameters. In fact MCDA has been criticised for not being an exact procedure and therefore it is very important to test how robust the results are. A second reason is that the SA of MCDA procedure enables the data and the decision-making problem to be explored at greater depths. This provides insight into the nature of the decision problem unravels its complexities and may even provide recommendations for future analysis. Furthermore, a SA may be carried out in order to deal with the uncertainty in estimation of some of the input figures. The decision-makers may not be able to derive at a set of criteria weights and may provide a range of weightings or the decision involves various groups of stakeholders with different preferences towards the weights which can be analysed. In the same line, the impacts of various options under different criteria may fall within a statistically estimated range that can be incorporated into the analysis (Proctor and Qureshi [15]).

As MCDA includes a comprehensive process involving a rich interaction between human judgement, data analysis and mathematical/computational processes the need for SA is clear. Errors and unintentional biases can enter at any of these stages, and it is the process as a whole that needs to be robust. Perhaps some of the key points at which such errors and biases may intrude

would be the following (Stewart [16]): external uncertainties, internal uncertainties, choice of model and identification of criteria and alternatives.

Figure 1 illustrates some of these key points at the different levels of uncertainty of a MCDA process. The two top levels of uncertainty are subjected to SA. Here internal uncertainties, choice of model and the identification of criteria and alternatives can be examined. Uncertainties on the environment (external) are often not analysed in SA.
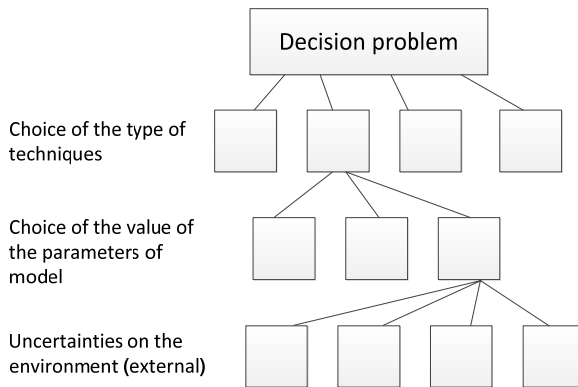

Figure 2:   Different levels of uncertainty (Vincke [17]).

The proposed methodology in this paper for examining the robustness of MCDA applied to the decision problem operates on different levels of the decision problem. The choice of MCDA model may affect the overall robustness of the analysis. This is with respect to the determination of weight and the method used for assessment of the alternatives performance on different attributes. In this paper a restriction to examining MCDA methodologies using pair-wise comparisons is made. This restriction is a result of experience conducting decision support. This paper will examine how the choice of assessment methodology can affect the overall outcome of the analysis. This is done be applying the same decision problem to the AHP and REMBRANDT methods (respectively additive and multiplicative methods using pair-wise comparisons) and the examining how robust the results are for changes in the attributes weights. If people are unsure of the weighting methods/principle a robustness measure of changes in the weights are more relevant – with regard to the uncertainty in the weighting. But even if people are sure of the methods a robustness analysis could be useful to deal with changes in or different preferences. Elicitation of weights is usually a time consuming process and is often controversial, as well it is difficult to derive exact weights, and it is also difficult to determine consistent boundaries for the intervals within actual weights are located. In such circumstances, ordinal ranking could be a reasonable compromise that uses input of consistent information and often provides output rank order of alternatives similar to the rank order based upon cardinal information (Shepetukha and Olson [18]). In this paper the ordinal ranking

method Rank Order Distribution (ROD) for assessing criteria weights is used (Roberts and Goodwin [19]).

The case studies in this paper are all decision problems where the different stakeholders had agreed upon which criteria to include in the MCDA or in other words they had all agreed upon the structure of the decision problem. Thereby the difference among the stakeholders can be described as differences in assessment of the different alternatives with respect to each criterion and differences in weights. The latter is described in the classical work of Edwards and Newman [20], where they have found that resolving conflicts about weights are easier to interpret and discus than conflicts about the structure of the decision problem, e.g. which criterion to include or not.

Three different classes of simulation are used to assess the robustness of the decision: (1) Random ROD weights without importance order – simulation where criteria are assigned a random ROD weight. Two criteria can have equal weighting. (2) Random order weights – the importance order of the criteria is preserved, but the weights are randomly generated. (3) Weight interval – identifying critical criteria where a change in the weight will result in different ranking of the alternatives. The three classes can be interpreted as different levels of knowledge. With random weights there exists no knowledge about the relative importance of the weights and is an extreme case of SA. For Random order weights the rank order of the criteria is maintained, but the weights are otherwise generated at random. Lastly, for weight interval it is assumed that the weights assessed by the stakeholders or decision-makers are valid, but with some uncertainty and the simulation can identify the critical criteria where a change of weight can result in a different ranking of the alternatives. This information can be useful, especially within the public decision domain where there can be a need to justify the decision for the public. An indication of how critical the criteria are if the weights are changed can defuse a potentially lengthy debate about the setting of the weights (Butler *et al.* [12]).

In order for using SA active and directly in the decision process with the stakeholders there is a need for specific properties of the SA. First the SA needs to be a transparent methodology in order for getting acceptance and thereby the necessary attention from the stakeholders. Second, the SA has to able to produce measures for robustness fairly fast, both with respect to the set-up and computation time.

3   Results

Two case studies are used for illustration. Both of the case studies are larger strategic transport projects which are highly complex and also are of great interest for the general public. The first case is considering the alternatives for a new fixed link between the two cities Helsingør (Elsinore) in Denmark and Helsingborg in Sweden (referred to as the HH-connection). Here the decision problem is to decide between three alternatives supporting different modes of transport. The second case is an examination of alignment of a high speed railway – Östlänken in Sweden. The decision problem is to decide on which of

four different alternatives for the alignment on a smaller segment of the project. Common for the two case studies is that in both cases a decision conference gathering some of the stakeholders was conducted in order to get input to the decision analysis.

3.1 Case study: HH connection

This case study is about an additional fixed link between Denmark and Sweden. In the autumn of 2011 three tunnel alternatives are considered as main alternatives for the HH-connection: Alt1 is a tunnel for rail, passenger trains only. Alt2 is tunnel for rail, passenger trains only and a tunnel for vehicles. The last alternative, Alt3, is tunnel for rail, passenger trains, a tunnel for vehicles and a tunnel for rail goods trains.

The first SA examined a random importance order of the criteria and the resulting total score for the alternatives. The results revealed that Alt1 will not be able to compete with the two other alternatives in any of the simulations. In Figure 3 the distributions of the simulated scores for Alt2 and Alt3 are shown.
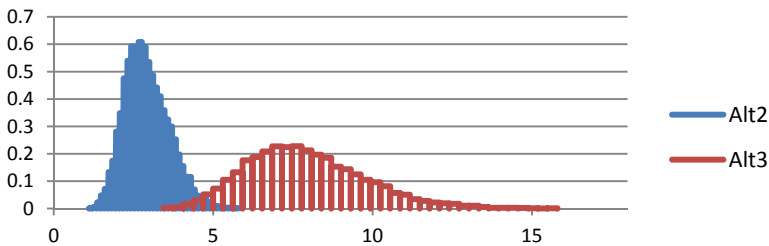

Figure 3:   Simulation with ROD weights based on random importance order of the criteria.

It can be seen from Figure 4 that Alt3 is most likely to the most preferred alternative and only in a small subset of the simulation results Alt2 can compete with Alt3. If the importance order of the criteria assessed by the stakeholders are preserved the simulation results will show the same tendency as with random weights, however the spread of the scores will be smaller (Figure 4).

In order to pinpoint which criteria are critical for the ranking of the alternatives the weight stability SA was performed. The results shown in Table 1 indicate that the 'Socio-economic robustness' and 'Impact on towns and land-use' criteria are critical for the ranking of the alternatives. The column 'Initial' shows the criteria weights used, based on the rank order of the criteria.

As it can be seen both the REMBRANDT and AHP MCDA methodologies have indicated the same critical criteria. However, AHP is much more sensitive to a change in the criterion 'impact on towns and land-use'.
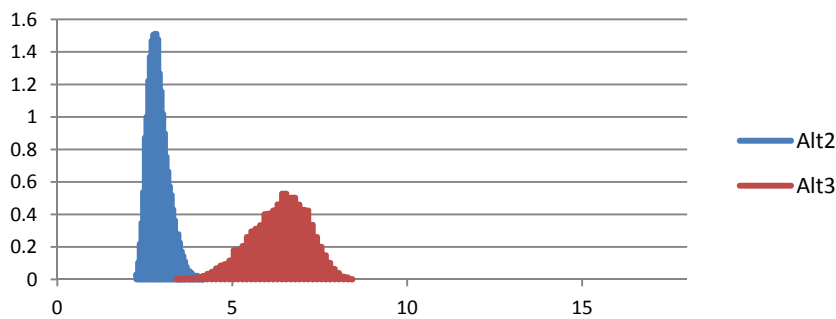
Figure 4:   Simulation with random generated weights assigned to the criteria preserving the importance order of the criteria.

Table 1:   Weight stability intervals of HH case.

		REMBRANDT		AHP	
Criterion	Initial	Lower	Upper	Lower	Upper
Socio-economic robustness	0.24	-	0.47	-	0.48
Improvement for passenger cars and public transport	0.19	-	-	-	-
Impact on towns and land-use	0.04	-	0.49	-	0.15
Impact on regional economics	0.30	-	-	-	-
Impacts on flexibility in logistics	0.09	-	-	-	-
Contribution to the EU green corridors	0.14	-	-	-	-

3.2 Case study: Östlänken

A new high speed railway line is planned in Sweden connecting Stockholm with Göteborg and Malmö. This case study concentrates on the link between Norrköping and Backeby. 4 different corridors have been identified and the decision problem is to choose between one of these 4 corridors.

The first SA examined a random importance order of the criteria and the resulting total score for the alternatives. The simulation results reveal that the final ranking of alternatives is sensitive to the criteria weights (Figure 5). Alt1 is in most of the simulation the lowest ranked alternative and in none of the simulations the highest ranked. For Alt2 and Alt3 the simulation results are more ambiguous. Alt2 tends to be higher ranked by Alt3, but only vaguely. Alt4 are ranked as third alternative in most of the simulations, but a more detailed analysis of the results revealed that with some special weight combinations the alternative is the highest ranked. However, this was just in a minor subset of the simulations and only if the REMBRANDT methodology was applied.
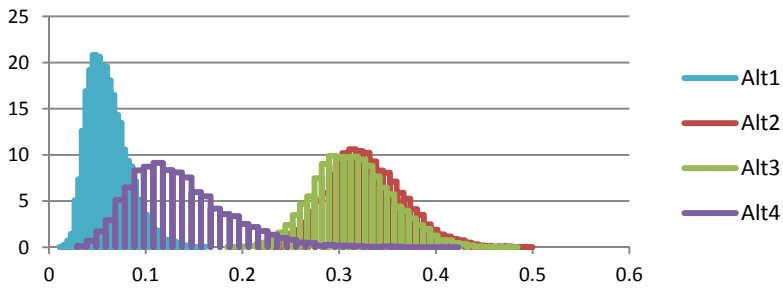
Figure 5:   Simulation with ROD weights based on random importance order of the criteria.

In Figure 6 the simulation results are illustrated when the importance order is preserved. From this figure some of the same conclusions can be made as with Figure 5, however the spread of the scores are smaller for each of the alternatives. The overall conclusions from these two first SA is that it is most likely to be alternative 2 and 3 that will get the highest score and it is not possible to say which of them is the most preferred.
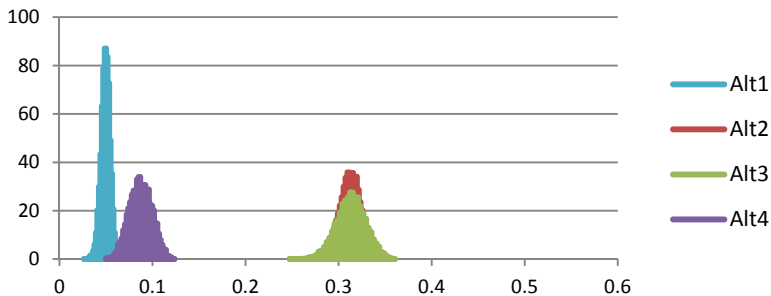

Figure 6:   Simulation with random generated weights assigned to the criteria preserving the importance order of the criteria.

For the Östlänken case the weight stability intervals for the criteria has been calculated and are shown in Table 2. The column 'Initial' shows the criteria weights used, based on the rank order of the criteria. From the table it can be seen that a change in several of the criteria weights can have influence on the ranking of the alternatives. The critical criteria weight for 'City and scenery impression' and 'Health' are relative close to the initial weights – especially if the REMBRANDT methodology is applied. Also a minor lowering of the weight on the criterion 'Natural environment' will result in a different ranking. A major reason for the large number of critical criteria in this case is that Alt2 and Alt3 are assessed to be very close to each other. So the critical criteria reflect that changes in these weights will result in a shift between these two alternatives as being the most preferred one.

Table 2:   Weight stability intervals.

		REMBRANDT		AHP	
Criterion	Initial	Upper	Lower	Upper	Lower
Socio-economic	0.21	0.37	-	0.51	-
City and scenery impression	0.18	0.27	-	0.49	-
Cultural environment	0.16	0.53	0.06	0.33	-
Natural environment	0.13	-	0.09	-	0.05
Health	0.11	0.16	-	0.22	-
Natural ressources	0.09	-	-	0.89	-
Risk and safety	0.06	0.22	-	0.21	-
Recreation and outdoor life	0.04	-	-	-	-
Building time	0.02	0.40	-	0.21	-

With regard to the two MCDA techniques, this case study implies that the REMBRANDT method is more sensitive to changes in the weights than the AHP.

4   Discussion and conclusion

This paper has illustrated a combination of simulation and rank order SA technique in order to provide the decision-makers with a clear picture of the robustness of the decision at stake. The technique presented is transparent and possible to use actively during the decision-making process and has been illustrated by two case studies.

The two case studies indicated that it is possible to produce a transparent and valid analysis of the robustness of a MCDA process to be used actively by the stakeholders or decision-makers. By working on different levels of the decision problem and thereby seek to simplify the SA, a clear and informative analysis can be conducted.

For the two case studies the random ROD weights approach gives information about how sensitive the decision analysis is for other ranking of the criteria after importance. This is useful if not all of the stakeholders or decision-makers can agree on one particular ranking order of criteria. Furthermore the random weight approach, where the ranking order is preserved, shows the effect of using other weights than the applied ROD weights. This can be used to comply with critique of the ROD methodology. By examining the weight stability intervals of the criteria, critical criteria can be found. For the HH case two critical criteria was found, however a rather large change in the weights from the initial assessed weights have to be applied in order to get another ranking order of the alternatives. For the Östlänken case several critical criteria were found, this is a result of the two close competing alternatives. The results from the last approach can be used as an indication of which criteria to examine

further in the decision problem, in order to provide a solid foundation for at decision.

Two different types of aggregation methods have been applied to the two case studies, the AHP and the REMBRANDT method. The results from the two case studies cannot give any clear conclusion on which method is more sensitive than the other with respect to the criteria weights. However, this only implies that there could be a difference in the resulting ranking of alternatives between these two methods and one has to be aware of this.

It is important also to emphasis on possible conflicts about the structure of the decision problem, and not only on the criteria weights. However, it has been the purpose of this paper to compare and produce some recommendations for SA to be used actively during a decision process with regard to the uncertainties and/or difference in preferences expressed by the criteria weights.

The tools for analysing the robustness of the MCDA must be able to communicate the complexity and associated uncertainties of the decisions and to allow for broad stakeholder participation while integrating different aspects of the situation involved. We should employ methods that are theoretically sound and, at the same time, practical and effective decision support tools.

References

[1] W.E. Walker, "Policy analysis: a systematic approach to supporting policymaking in the public sector," *Journal of Multicriteria Decision Analysis*, vol. 9, no. 1, pp. 11-27.

[2] J. Ananda and G. Herath, "A critical review of multi-criteria decision making methods with special reference to forest management and planning," *Ecol. Econ.*, vol. 68, no. 10, pp. 2535-2548.

[3] S. Damart and B. Roy, "The uses of cost-benefit analysis in public transportation decision-making in France," *Transp. Policy*, vol. 16, no. 4, pp. 200-212.

[4] G. Munda, "Social multi-criteria evaluation: Methodological foundations and operational consequences," *Eur. J. Oper. Res.*, vol. 158, no. 3, pp. 662-677.

[5] S.O. Funtowicz and J.R. Ravetz, "Science for the post-normal age," *Futures*, vol. 25, no. 7, pp. 739-755.

[6] A.C. Petersen, A. Cath, M. Hage, E. Kunseler and d.S. van, "Post-Normal Science in Practice at the Netherlands Environmental Assessment Agency," *Science, Technology, & Human Values*, vol. 36, no. 3, pp. 362-388.

[7] E. Triantaphyllou and A. Sánchez, "A sensitivity analysis approach for some deterministic multi-criteria decision-making methods," *Decision Sciences*, vol. 28, no. 1, pp. 151-185.

[8] A.A. Salo and R.P. Hamalainen, "Preference Assessment by Imprecise Ratio Statements," *Oper. Res.*, vol. 40, no. 6, Nov, pp. 1053.

[9] M.P. Strager and R.S. Rosenberger, "Incorporating stakeholder preferences for land conservation: Weights and measures in spatial MCA," *Ecol. Econ.*, vol. 58, no. 1, pp. 79-92.

[10] A. Jessop, "Using imprecise estimates for weights," *J. Oper. Res. Soc.*, vol. 62, no. 6, pp. 1048-1055.

[11] D.R. Insua and S. French, "A framework for sensitivity analysis in discrete multi-objective decision-making," *Eur. J. Oper. Res.*, vol. 54, no. 2, pp. 176-190.

[12] J. Butler, J. Jia and J. Dyer, "Simulation techniques for the sensitivity analysis of multi-criteria decision models," *Eur. J. Oper. Res.*, vol. 103, no. 3, 12/16, pp. 531-546.

[13] J. Butler and D.L. Olson, "Comparison of centroid and simulation approaches for selection sensitivity analysis," *Journal of Multi-Criteria Decision Analysis*, vol. 8, no. 3, pp. 146-161.

[14] S. French, J. Maule and N. Papamichail, "Decision behaviour, analysis and support," 2009.

[15] W. Proctor and M. Qureshi, "Multi-criteria Evaluation Revisited," 11-13 December 2005.

[16] T.J. Stewart, "Robustness Analysis and MCDA".

[17] P. Vincke, "About Robustness Analysis".

[18] Y. Shepetukha and D.L. Olson, "Comparative analysis of multiattribute techniques based on cardinal and ordinal inputs," *Math. Comput. Model.*, vol. 34, no. 1–2, 7, pp. 229-241.

[19] R. Roberts and P. Goodwin, "Weight approximations in multi-attribute decision models," *Journal of Multi Criteria Decision Analysis*, vol. 11, pp. 291-304.

[20] W. Edwards and R. Newman J., "Multiattribute Evaluation," 1982.