ProMCDA: A Python package for Probabilistic Multi-Criteria Decision Analysis

Flaminia Catalli [1*] **and Matteo Spada** [2*]

1 wetransform GmbH, Germany **2** Zürich University of Applied Sciences, School of Engineering, INE Institute of Sustainable Development, Switzerland * These authors contributed equally.

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Gabriela Alessio Robles

Reviewers:

- @limengbinggz
- @B3J4y

Submitted: 28 November 2023
Published: unpublished

Summary

Multi-Criteria Decision Analysis (MCDA) is a formal process to assist decision makers (DMs) in structuring their decision problems and to provide them with tools and methods leading to recommendations on the decisions at stake (Roy (1996)). The recommendations are based on a comprehensive identification of the alternatives considered and the selection of criteria/subcriteria/etc. to evaluate them, which are aggregated taking into account the preferences of the DMs (Bouyssou et al. (2006)). In the literature, there is a wide range of MCDA methods used to integrate information and either classify alternatives into preference classes or rank them from best to worst (Cinelli et al. (2022)). In the context of ranking and benchmarking alternatives across complex concepts, composite indicators (CIs) are the most widely used synthetic measures (Greco et al. (2019)). Indeed, they have been applied, for example, in the context of environmental quality (Oțoiu & Grădinaru (2018)), resilience of energy supply (Gasser et al. (2020)), sustainability (Volkart et al. (2016)), global competitiveness (Klaus Schwab (2018)), etc. However, the uncertainty of the criteria, the effect of tuning the weights relative to them, and the choice of methods (normalization/aggregation) to construct CIs, have been shown to influence the final ranking of alternatives (e.g. Cinelli et al. (2020)).

The `ProMCDA` Python module proposed here allows a DM to explore the sensitivity and robustness of the CIs results in a user-friendly way. In other words, it allows the user to assess either the sensitivity related to the choice of normalization and/or aggregation method, but also to account for uncertainty in the criteria and weights.

Statement of need

There are already dedicated tools for estimating CIs in the literature. In *R*, there is an existing package called `COINr`, which allows the user to develop CIs by including all common operations, from criteria selection, data treatment, normalization and aggregation, and sensitivity analysis (Becker et al. (2022)). There are also other packages in R, such as `compind`, that focus on weighting and aggregation (Fusco et al. (2018)). In *MATLAB*, there are some packages dedicated to specific parts of CI development, such as the `CIAO` tool (Lindén et al. (2021)). The Python module `Decisi-o-Rama` (Chacon-Hurtado & Scholten (2021)) focuses on the implementation of the Multi-Attribute Utility Theory (MAUT) to normalize criteria, considering a hierarchical criteria structure and uncertain criteria, and to aggregate the results using different aggregation methods. Finally, the web tool called `MCDA Index Tool` allows sensitivity analysis based on different combinations of normalization functions and aggregation methods (MCDA Index Tool).

`ProMCDA` is a Python module for performing CIs MCDA considering a full probabilistic approach.

The tool provides sensitivity and robustness analysis of the ranking results. The sensitivity of the MCDA scores is caused by the different pairs of normalization/aggregation functions (Cinelli et al. (2020)) that can be used in the evaluation process. The uncertainty is instead caused by either the variability associated with the criteria values (Stewart & Durbach (2016)) or the randomness that may be associated with their weights (Lahdelma et al. (1998)). `ProMCDA` is unique in combining all these different sources of variability and providing a systematic analysis.

The tool is designed to be used by both researchers and practitioners in operations research. The approach has a wide range of potential applications, ranging from sustainability to healthcare and risk assessment, to name but a few. `ProMCDA` has been developed as a core methodology for the development of a decision support system for forest management (FutureForest). However, the tool is generic and can be used in any other domain involving multi-criteria decision-making.

Overview

`ProMCDA` is a module consisting of a set of functions that allow CIs to be constructed considering the uncertainty associated with the criteria, the weights and the combination of normalization/aggregation methods. The evaluation process behind `ProMCDA` is based on two main steps of data manipulation:

- data normalisation, to work with data values on the same scale;
- data aggregation, to estimate a single composite indicator from all criteria.

`ProMCDA` receives all the necessary input information via a configuration file in JSON format (for more details see the README). The alternatives are represented in an input matrix (in CSV file format) as rows and described by the different values of the criteria in the columns. The sensitivity analysis is performed by comparing the different scores associated with the alternatives, which are obtained by using different combinations of normalization and aggregation functions. `ProMCDA` implements 4 different normalization and 4 different aggregation functions, as described in Table 1 and Table 2 respectively. However, the user can decide to run `ProMCDA` with a specific pair of normalization and aggregation functions, and thus switching off the sensitivity analysis.

Normalization methods		Formula	Descript
Linear scale	**Min-max**	$N_{ia} = \frac{x_{ic} - \min(x_i)}{\max(x_i) - \min(x_i)}$	It applies transform rescale th specified (typically (
	Standardization (z-score)	$N_{ia} = \frac{x_{ia} - x_{ia=\bar{a}}}{\sigma_{ia=\bar{a}}}$	It applies transform mean of standard of 1.
Ratio scale	**Target**	$N_{ia} = \frac{x_{ia}}{\max(x_i)}$	It norma upper limi
Ordinal	**Rank**	$N_{ia} = rank(x_{ia})$	The data ranked b their relat

Legend

N_{ia}: the normalized value of indicator *i* for alternative *a*.
x_{ia}: the value of indicator i for alternative *a*.
$x_{ia=\bar{a}}$the average value of indicator *i* across all alternatives.
$\sigma_{ia=\bar{a}}$: the standard deviation of indicator *i* across all alternatives.
$min(x_i)$: the minimum value of indicator *i* across all alternatives.
$max(x_i)$: the maximum value of indicator *i* across all alternatives.

Table 1: Normalization functions used in `ProMCDA`*.*

Table 2: Aggregation functions used in `ProMCDA`*.The sum of the weights is normalized to 1 as in*

Aggregation methods	Formula	Level of compensation	Comments
Additive (weighted arithmetic mean)	$score_a = \sum_{i=1}^{n} N_{ia}\, w_i$	Full	Most common a used. It is a line amplifies the ef values. Comm situations whe considered equa
Geometric (weighted geometric mean)	$score_a = \prod_{i=1}^{n} {N_{ia}}_i^{\,w_i}$	Partial	The indicators larger than 0. combination. Th variable's value i to its magnitude contribution o depends on th involved. It ampl variables with method is co situations where joint effect of var
Harmonic	$score_a = \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} \frac{w_i}{N_{ia}}}$	Partial (less than Geometric)	The indicators va be larger than 0 combination. Th value is not p magnitude, an contribution o depends on th involved. Insens values. It is p situations where considered mo when dealing wi
Minimum	$N_{ia} = min(N_{1a}, N_{2a}, \ldots, N_{na})$	None	The worst per equals the final s DM is intereste driven by the indicator.

Legend

$score_a$: the composite score for alternative *a*.
n: the number of indicators.
w_i: the weight of indicator *i*.
N_{ia}: the normalized value of indicator *i* for alternative *a*.

Langhans et al. (2014).

The user can also decide to run ProMCDA with or without a robustness analysis. The robustness analysis is triggered by adding randomness to either the weights or the criteria. This means that either the weights or the criteria values are randomly sampled using a Monte Carlo method. In ProMCDA randomness is not allowed for both weights and criteria in order to make the results as transparent as possible. In fact, mixing uncertainty from both weights and criteria would lead to a lack of distinction between the effect of one or the other. Randomness in the weights can be applied to one weight at a time or to all weights at the same time. In the first case, the aim is to be able to analyse the effect of each individual criteria on the scores; in the second case, it is to have an overview of the uncertainty associated with all the weights. In both cases, by default, the weights are sampled from a uniform distribution [0-1]. On the other hand, if the user decides to analyse the robustness of the criteria, he/she has to provide the parameters defining the marginal distribution (i.e. a probability density function, pdf) that best describes the criteria, rather than the criteria values. This means that if a criterion is characterized by a pdf described by 2 parameters, two columns should be allocated in the input CSV file for it. In ProMCDA 4 different pdfs describing the criteria uncertainty are considered:

- *uniform*, which is described by 2 parameters, i.e., minimum and maximum
- *normal*, which is described by 2 parameters, i.e., mean and standard deviation
- *lognormal*, which is described by 2 parameters, i.e., log(mean) and log(standard deviation)
- *Poisson*, which is described by 1 parameter, i.e., the rate.

Once the pdf for each criterion is selected and the input parameters are in place in the input CSV file, ProMCDA randomly samples n-values of each criterion per alternative from the given pdf and assesses the score and ranking of the alternatives, by considering robustness at the criteria level. The number of samples is given in the configuration file by the user.

Once the pdfs for each criterion are selected and the input parameters are in the input CSV file, `ProMCDA` randomly samples n-values of each criterion per alternative from the given pdf to evaluate the score and ranking of the alternatives, taking into account robustness at the criteria level.

Finally, in all possible cases (i.e. a simple MCDA; MCDA with sensitivity analysis for the different normalization/aggregation functions used; MCDA with robustness investigation related either to randomness on the weights or on the indicators), `ProMCDA` will output a CSV file with the scores/average scores and their plots. For a quick overview of the functionality of `ProMCDA`, refer to Table 3. For more details, refer to the README.

Possible usages of ProMCDA	Specifications
Simple MCDA No sensitivity nor robustness analysis is performed.	The specific pair normaliza /aggregation to be used for evaluation of the alternatives.
Sensitivity analysis Focus is on the role of the normalization and aggregation functions.	All normalization and aggregation p are used for the evaluation of alternatives.
Robustness analysis of one weight at time Focus is on the role of one indicator and its relative weight at time.	One single weight at time is samp from the uniform distribution [0,1].
Robustness analysis of all weights Focus is on the role of the weights.	All weights are sampled from uniform distribution [0,1].
Robustness analysis of the indicators Focus is on the role of the uncertainty of the indicators.	All indicators, whose values distributed as a non-exact pdf, randomly sampled. *ProMCDA* needs values for each indicator per alterna to build N random input-matrices.

Table 3: Overview on the functionalities of ProMCDA.

Acknowledgements

Flaminia Catalli was supported by the Future Forest II project funded by the Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (Germany) grant Nr. 67KI21002A. The authors would like to thank Kapil Agnihotri for thorough code revisions, Thorsten Reitz and the whole Future Forest II team for productive discussions on a problem for which we have found a robust and transparent solution over time.

References

Becker, W., Caperna, G., Sorbo, M. D., Norlén, H., Papadimitriou, E., & Saisana, M. (2022). COINr: An R package for developing compositeindicators. *Journal of Open Source Software*, *7*(78), 4567. https://doi.org/10.21105/joss.04567

Bouyssou, D., Marchant, T., Pirlot, M., Tsoukiàs, A., & Vincke, P. (Eds.). (2006). Problem Formulation and Structuring: The Decision Aiding Process. In *Evaluation and Decision Models with Multiple Criteria: Stepping stones for the analyst* (pp. 19–65). Springer US. https://doi.org/10.1007/0-387-31099-1_2

Chacon-Hurtado, J. C., & Scholten, L. (2021). Decisi-o-rama: An open-source Python library for multi-attribute value/utility decision analysis. *Environmental Modelling & Software*, *135*, 104890. https://doi.org/10.1016/j.envsoft.2020.104890

Cinelli, M., Burgherr, P., Kadziński, M., & Słowiński, R. (2022). Proper and improper uses of MCDA methods in energy systems analysis. *Decision Support Systems*, *163*, 113848. https://doi.org/10.1016/j.dss.2022.113848

Cinelli, M., Spada, M., Kim, W., Zhang, Y., & Burgherr, P. (2020). MCDA Index Tool: An interactive software to develop indices and rankings. *Environment Systems and Decisions*. https://doi.org/10.1007/s10669-020-09784-x

Fusco, E., Vidoli, F., & Sahoo, B. K. (2018). Spatial heterogeneity in composite indicator: A methodological proposal. *Omega*, *77*, 1–14. https://doi.org/10.1016/j.omega.2017.04.007

Gasser, P., Suter, J., Cinelli, M., Spada, M., Burgherr, P., Hirschberg, S., Kadziński, M., & Stojadinovic, B. (2020). Comprehensive resilience assessment of electricity supply security for 140 countries. *Ecological Indicators*, *110*. https://doi.org/10.1016/j.ecolind.2019.105731

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the Methodological Framework of Composite Indices: A Review of the Issues of Weighting, Aggregation, and Robustness. *Social Indicators Research*, *141*(1), 61–94. https://doi.org/10.1007/s11205-017-1832-9

Klaus Schwab. (2018). *The global competitiveness report 2018*. https://apo.org.au/node/262576

Lahdelma, R., Hokkanen, J., & Salminen, P. (1998). SMAA - Stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, *106*(1), 137–143. https://doi.org/10.1016/S0377-2217(97)00163-X

Langhans, S. D., Reichert, P., & Schuwirth, N. (2014). The method matters: A guide for indicator aggregation in ecological assessments. *Ecological Indicators*, *45*, 494–507. https://doi.org/10.1016/j.ecolind.2014.05.014

Lindén, D., Cinelli, M., Spada, M., Becker, W., Gasser, P., & Burgherr, P. (2021). A framework based on statistical analysis and stakeholders' preferences to inform weighting in composite indicators. *Environmental Modelling & Software*. https://doi.org/10.1016/j.envsoft.2021.105208

Oţoiu, A., & Grădinaru, G. (2018). Proposing a composite environmental index to account for the actual state and changes in environmental dimensions, as a critique to EPI. *Ecological Indicators*, *93*, 1209–1221. https://doi.org/10.1016/j.ecolind.2018.06.009

Roy, B. (1996). Decision Aiding: Major Actors and the Role of Models. In B. Roy (Ed.), *Multicriteria Methodology for Decision Aiding* (pp. 7–17). Springer US. https://doi.org/10.1007/978-1-4757-2500-1_2

Stewart, T. J., & Durbach, I. (2016). Dealing with Uncertainties in MCDA. In S. Greco, M. Ehrgott, & J. R. Figueira (Eds.), *Multiple Criteria Decision Analysis: State of the Art Surveys* (pp. 467–496). Springer New York. https://doi.org/10.1007/978-1-4939-3094-4_12

Volkart, K., Bauer, C., Burgherr, P., Hirschberg, S., Schenler, W., & Spada, M. (2016). Interdisciplinary assessment of renewable, nuclear and fossil power generation with and without carbon capture and storage in view of the new Swiss energy policy. *International Journal of Greenhouse Gas Control*, *54*, 1–14. https://doi.org/10.1016/j.ijggc.2016.08.023