Evaluating the PROMETHEE II ranking quality

Boris Coquelet, Gilles Dejaegere, Yves De Smet

Université libre de Bruxelles
Ecole polytechnique de Bruxelles
CoDE-SMG

April 16, 2026

PROMETHEE - Quick overview

Developed by Brans and Vincke in 1982 Jean-Pierre Brans. *L'ingénierie de la décision: l'élaboration d'instruments d'aide à la décision*. Université Laval, Faculté des sciences de l'administration, 1982

Let us consider the set of alternatives $A = \{a_1, \cdots, a_n\}$ evaluated on the set of q criteria $F = \{f_1, \cdots, f_q\}$.

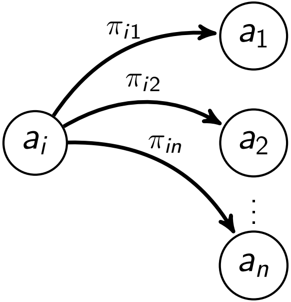
1. For each pair of alternatives, compute the pairwise preferences:

$$\pi_{ij} = \sum_{k=1}^{q} \omega_k \cdot \mathcal{F}_k(f_k(a_i) - f_k(a_j)) \tag{1}$$

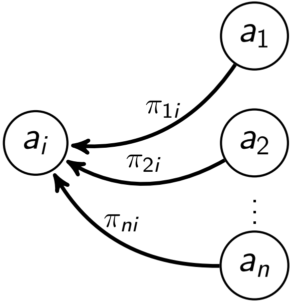
   Where ω_k is the weight of criteria k and $\mathcal{F}_k$ is a monotonically increasing function in $[0, 1]$.

2. For each alternative, compute the positive and negative outranking flow scores:

$$\phi_i^+ = \frac{1}{n-1} \sum_{a_j \in A, j \neq i} \pi_{ij} \qquad \phi_i^- = \frac{1}{n-1} \sum_{a_j \in A, j \neq i} \pi_{ji} \tag{2}$$



Preference of a_i over the rest



Preference of the rest over a_i

3. For each alternative, compute its net flow score:

$$\phi(a_i) = \phi^+(a_i) - \phi^-(a_i) \tag{3}$$

Research question

“ Is there a way to assess the quality of the results provided by the application of the PROMETHEE II method to a given dataset? “

Ex ante:

- Denis Bouyssou. “Ranking methods based on valued preference relations: A characterization of the net flow method”. In: *European Journal of Operational Research* 60.1 (1992), pp. 61–67
- Gilles Dejaegere, Mohamed Boujelben, and Yves De Smet. “An axiomatic characterization of Promethee II’s net flow scores based on a combination of direct comparisons and comparisons with third alternatives”. In: *Journal of Multi-Criteria Decision Analysis* (Mar. 2022)

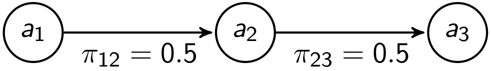
Ex post:

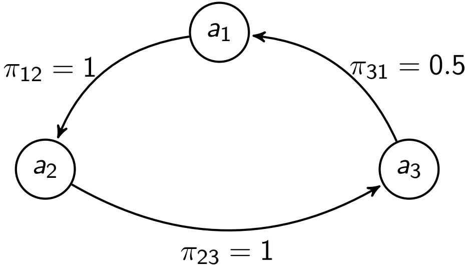
- Nguyen Anh Vu Doan and Yves De Smet. "An alternative weight sensitivity analysis for PROMETHEE II rankings". In: *Omega* 80 (2018), pp. 166–174
- Bertrand Mareschal. "Weight stability intervals in multicriteria decision aid". In: *European Journal of Operational Research* 33.1 (1988), pp. 54–64
- Alexandre Flachs and Yves De Smet. "Inverse optimization on the evaluations of alternatives in the Promethee II ranking method". In: *Omega* 136 (2025), p. 103325

Can we define a quality measure?

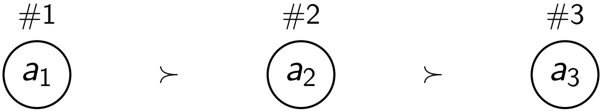
$\Rightarrow$ AHP has a similar idea with an inconsistency index Thomas L. Saaty. "The Analytic Hierarchy and Analytic Network Processes for the Measurement of Intangible Criteria and for Decision-Making". In: *Multiple Criteria Decision Analysis: State of the Art Surveys*. Ed. by Salvatore Greco, Matthias Ehrgott, and José Rui Figueira. Springer New York, 2016, pp. 363–419

Why? - Limit cases





In both cases, $\phi(a_1) = 0.25,\ \phi(a_2) = 0,\ \phi(a_3) = -0.25, .$ Thus:



Let us take a step back

When going from the pairwise preferences to the final ranking:

"PROMETHEE II ranks the alternatives to maximise the sum of all paths of at most length 2 starting from any alternative and ending on any alternative, worst in the ranking."

⇒ From the axiomatic developed by Dejaegere [4].
⇒ From the characterization of Bouyssou [1].

In mathematical form

PROMETHEE II ranking: $a_1 \succeq a_2 \succeq a_3 \succeq a_4 \succeq ... \succeq a_n$

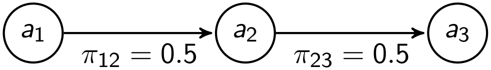
Maximize:

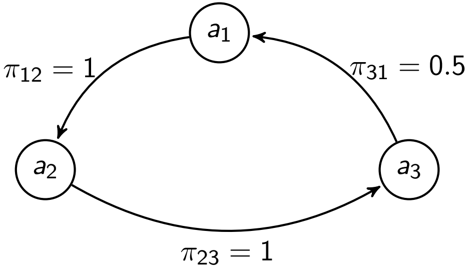
$$PL2 = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{k=1}^{n} (\pi_{ik} + \pi_{kj}) \tag{4}$$

Can be separated:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{k=1}^{n} (\pi_{ik} + \pi_{kj}) = (n-1) \left[\underbrace{\sum_{i=1}^{n} (n-i)\phi_i}_{dep.\ seq.} + \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{n} \pi_{ij}}_{indep.\ seq.} \right] \tag{5}$$

Equivalent problems





They are different in terms of: $\sum_{i=1}^{n} \sum_{j=1}^{n} \pi_{ij}$
Following the results of Bouyssou [1], removing cycles of the same value does not impact the final ranking.

$\Rightarrow$ Cycles artificially increases $\sum_{i=1}^{n} \sum_{j=1}^{n} \pi_{ij}$ and do not change the ranking

The quality index

- Cycles go against the transitive property of a consistent ranking;
- Cycles of a constant value do not impact $\phi_i \; \forall i = 1, ..., n$.

$\Rightarrow$ For a given set of ϕ_i, a set of pairwise preferences with no cycle is better represented by the ϕ_i than one with cycles.

- Cycles artificially increase $\sum_{i=1}^{n} \sum_{j=1}^{n} \pi_{ij}$;
- We can remove all cycles from the original set of pairwise preferences.

$\Rightarrow$ We can compare preference matrices on the sum.

However! Finding all cycles is exponentially time-consuming with the number of alternatives!

A simplified approach

Let us split the removal of cycles into two steps:

1. Removal of cycles of length 2;
2. Removal of the remaining cycles.

The first step is easy to compute. Let us define $\tilde{\pi}_{ij}$ the new pairwise preferences:

$$\tilde{\pi}_{ij} = \pi_{ij} - min(\pi_{ij}, \pi_{ji}) \quad and \quad \tilde{\pi}_{ji} = \pi_{ji} - min(\pi_{ij}, \pi_{ji}) \tag{6}$$

Which can be generalised:

$$\tilde{\pi}_{ij} = max(\pi_{ij} - \pi_{ji}, 0) \qquad \forall i, j \tag{7}$$

And we can define:

$$\tilde{\Pi} = \sum_{i=1}^{n} \sum_{j=1}^{n} max(\pi_{ij} - \pi_{ji}, 0) = \sum_{i=1}^{n} \sum_{j=i}^{n} |\pi_{ij} - \pi_{ji}| \tag{8}$$

First step vs second step

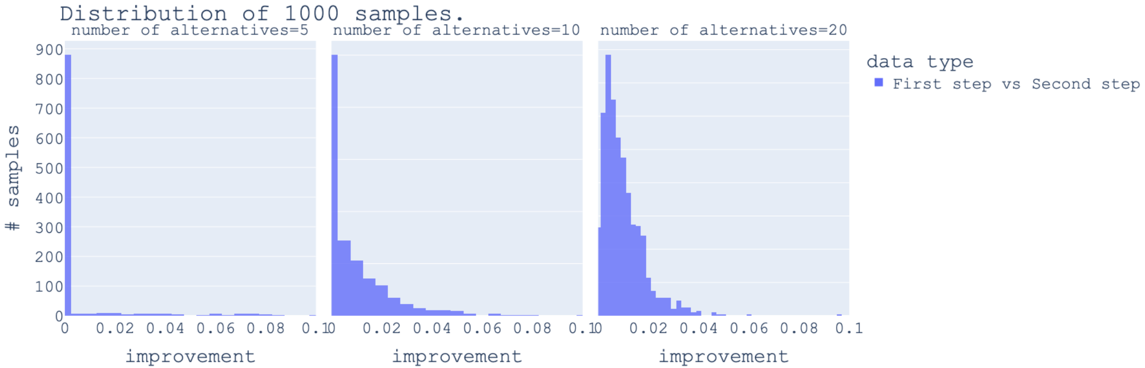


Figure: Improvement yield when applying the second step. Expressed as a percentage of $\tilde{\Pi}$ removed.

Quality index definition

Let:

$$\text{The initial pairwise preferences: } \Pi = \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_{ij} \tag{9}$$

$$\text{The modified pairwise preferences: } \tilde{\Pi} = \sum_{i=1}^{n} \sum_{j=i}^{n} |\pi_{ij} - \pi_{ji}| \tag{10}$$

We can define the following index:

$$\mathcal{I} = \frac{\Pi - \tilde{\Pi}}{\Pi} \qquad \mathcal{I} \in [0, 1] \tag{11}$$

$\Rightarrow \mathcal{I}$ is the percentage of the initial pairwise preferences that can be removed without modifying the net flow score.

Results

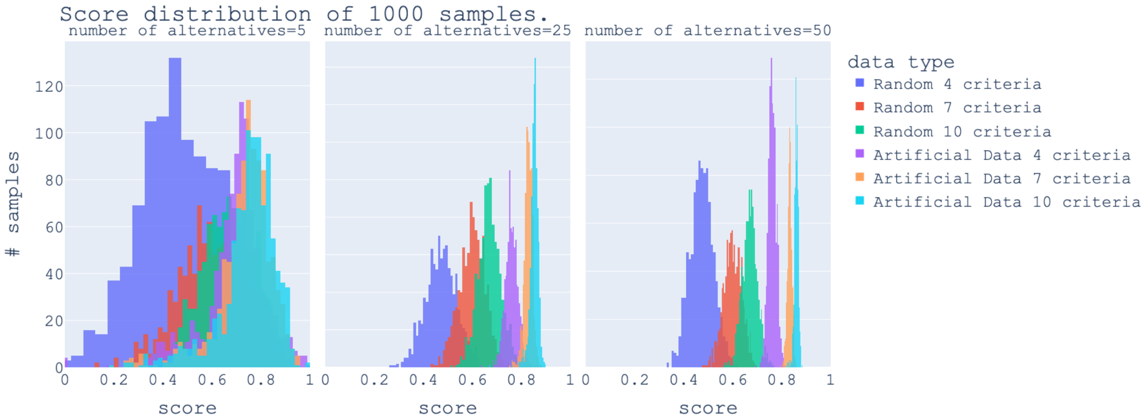


Figure: Score distribution of randomly generated data and extreme data.

Results (contd.)

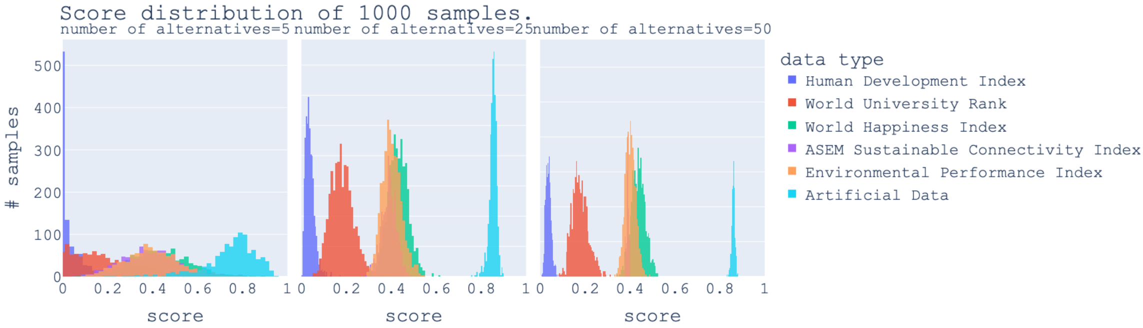


Figure: Score distribution of real data and extreme data.

A link with rank reversal?

Table: Comparison of the quality of the netflow score ranking and the proportion of pairs of alternatives subject to rank reversals in all datasets.

Datasets	$\mathcal{I}$	**% RR**
Human Development Index	0.03	0.23
World University Rank	0.18	0.30
World Happiness Index	0.46	0.65
ASEM Sustainable Connectivity Index	0.39	0.53
Environmental Performance Index	0.40	0.49
Artificial Data	0.86	1

Conclusion

- *"PROMETHEE II ranks the alternatives to maximise the sum of all paths of at most length 2 starting from any alternative and ending on any alternative, worst in the ranking."*
- Provided a quality index:

$$\mathcal{I} = \frac{\Pi - \tilde{\Pi}}{\Pi} \qquad \mathcal{I} \in [0, 1] \tag{12}$$

- Results on random, artificial and real datasets.
- Potential link with rank reversal.

Limitations:

- Limited the removal to cycles of length 2 (justified empirically).
- Dependence between the index values, the number of criteria and the parameters.
- No clear threshold of acceptance.

Currently exploring

- An interpretation of the value of the index;
- Removal of all cycles;
- The computation of net flow scores can be seen as a linear map applied to the pairwise preferences;
- An additional index for comparison/reinforcement;

Thank You!

Article available in open access:

Boris Coquelet, Gilles Dejaegere, and Yves De Smet. "Evaluating the Promethee ii Ranking Quality". In: *Algorithms* 18.10 (2025). ISSN: 1999-4893. DOI: 10.3390/a18100597. URL: https://www.mdpi.com/1999-4893/18/10/597