JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Empirical studies for comparing interactive multiobjective optimization methods with human participants

Misitano, G., Afsar, B., Silvennoinen, J., Ruiz A. B., Ruiz F., and Miettinen K.

EURO Conference 2024, Copenhagen, Denmark

1.7.2024

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Outline

- Motivation
- Background
- Experiment design
- Results
- Conclusions

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

About me

Giovanni Misitano

PhD

- Post-doctoral researcher from the University of Jyväskylä.
- Working in the Multiobjective Optimization Group led by prof. Kaisa Miettinen.
- My main research interest lies in the interface between decision makers and (interactive) multiobjective optimization methods.
- One of the main contributors and maintainers of the DESDEO software framework (more on this later).

Motivation

- Multiobjective optimization methods are tools to support decision-making.
- These tools can help decision makers find solutions to problems with multiple conflicting objective functions and no clear single optimum.
- Because of the large amount of optimal solutions, preferences can be utilized in some methods to guide the optimization process.
- Especially in so-called interactive multiobjective optimization methods, the decision maker can iteratively explore the set of optimal solutions and change their preferences, which allows them to learn about the available solutions and the feasibility of the preferences.

But how to compare interactive multiobjective optimization methods? How to choose the best interactive method for a decision maker?

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Background

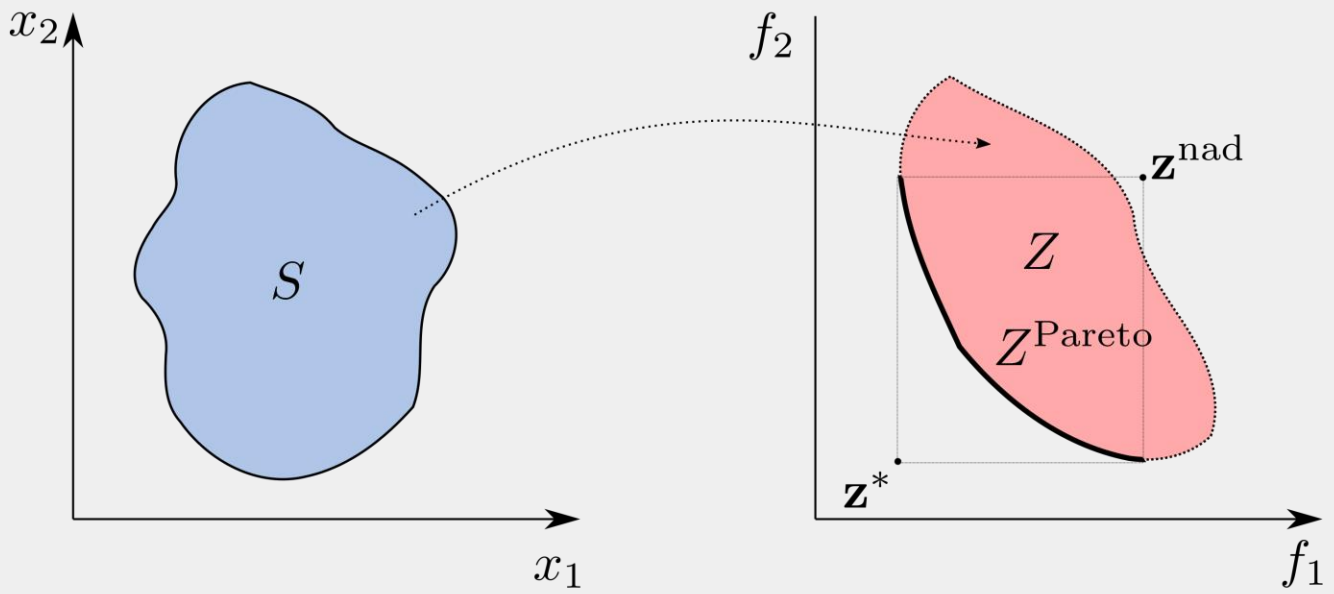
- Multiobjective optimization problem
- Multiobjective optimization methods
- Interactive methods
- Reference Point method (RPM)
- NIMBUS
- E-NAUTILUS
- Comparing interactive methods

Multiobjective optimization problem

Problem definition

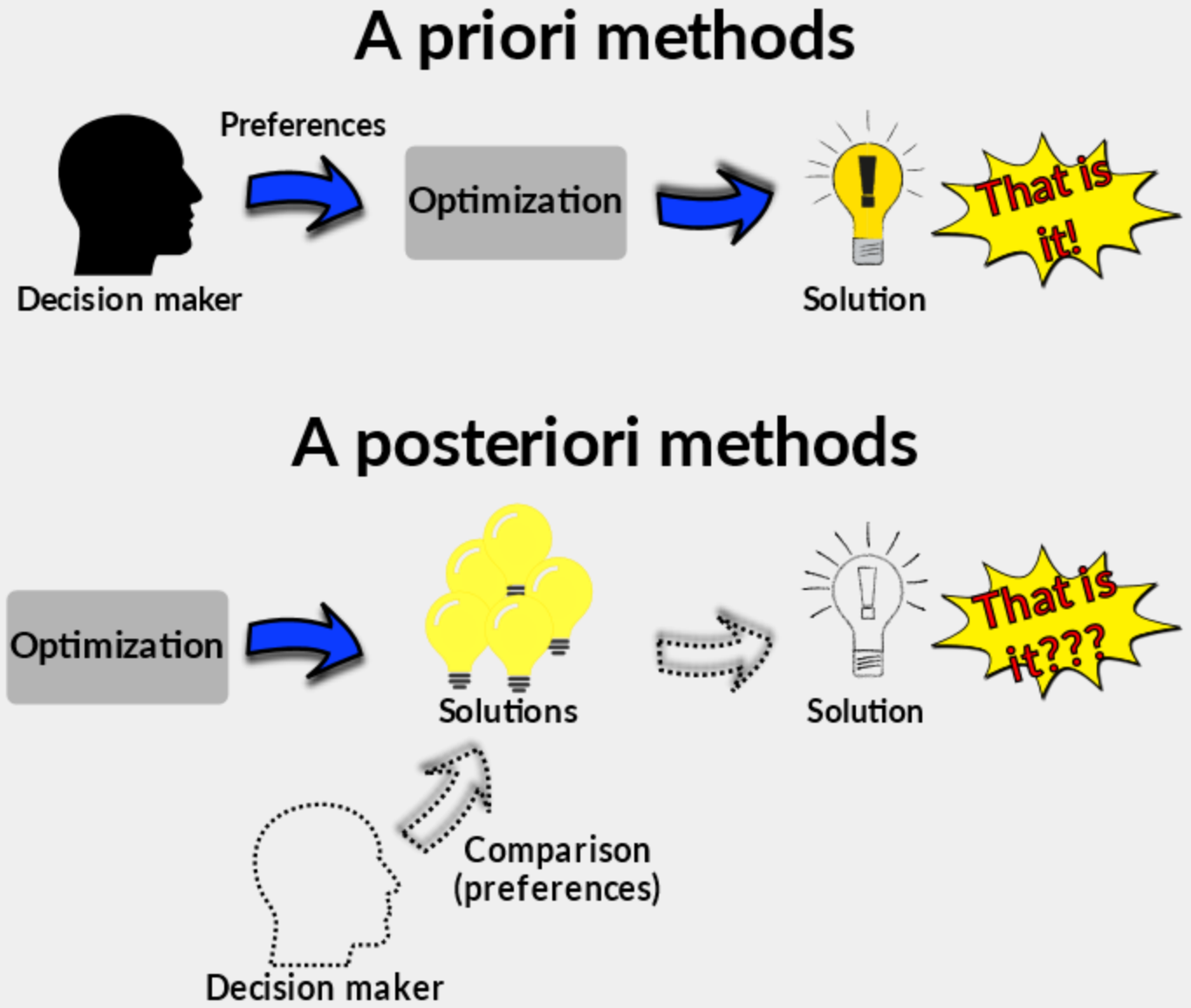
$$\begin{aligned} \text{minimize} \quad & F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})) \\ \text{s.t.} \quad & \mathbf{x} \in S \end{aligned}$$

Central concepts



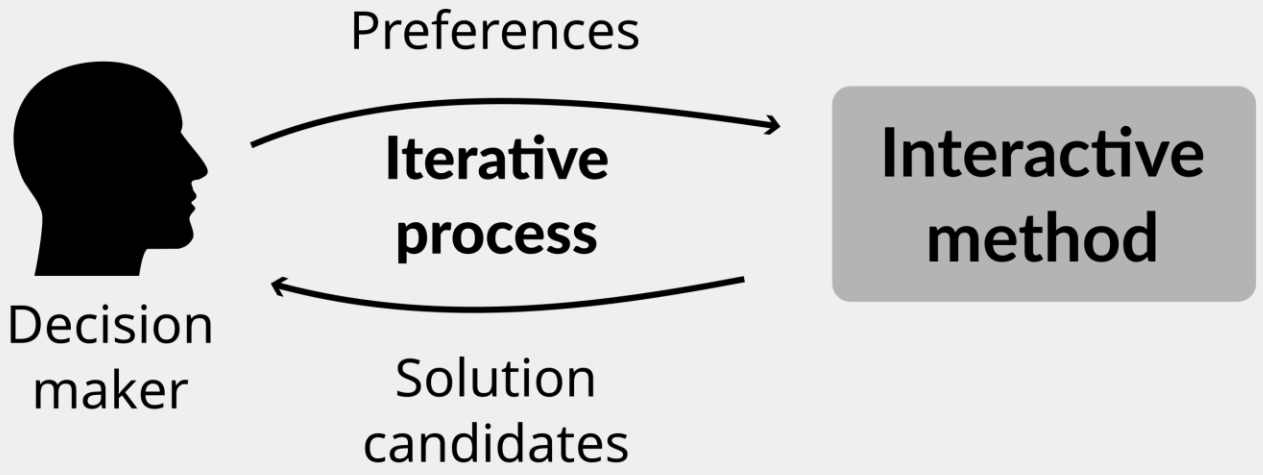
Miettinen, K. 1999. Nonlinear multiobjective optimization. Boston: Kluwer Academic Publishers.

Multiobjective optimization methods



Miettinen, K. 1999. Nonlinear multiobjective optimization. Boston: Kluwer Academic Publishers.
Sawaragi, Y., Nkayama, H., and Tanino, T. 1985. Theory of multiobjective optimization. Academic Press.

Interactive methods

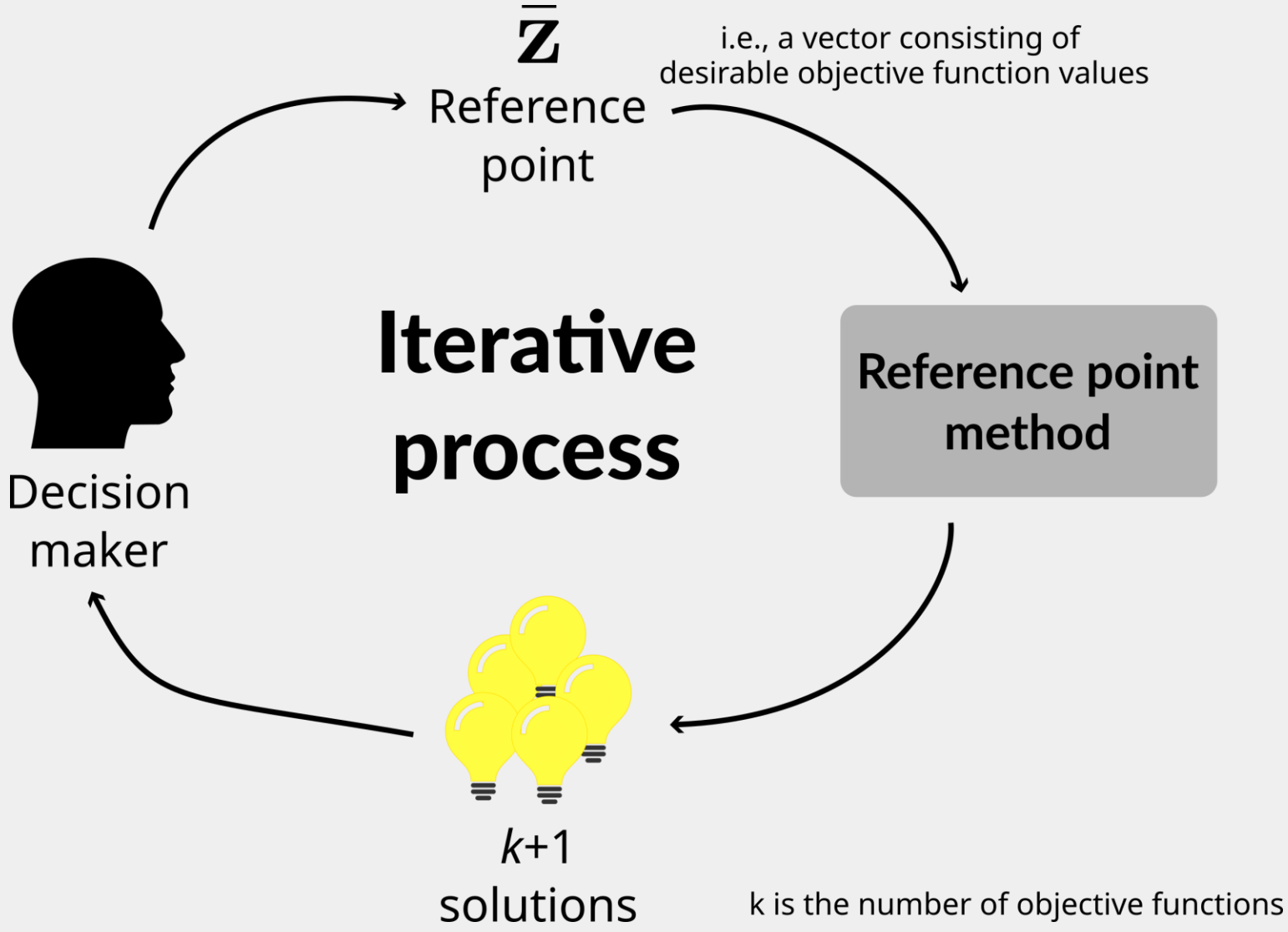


Exploration

Learning

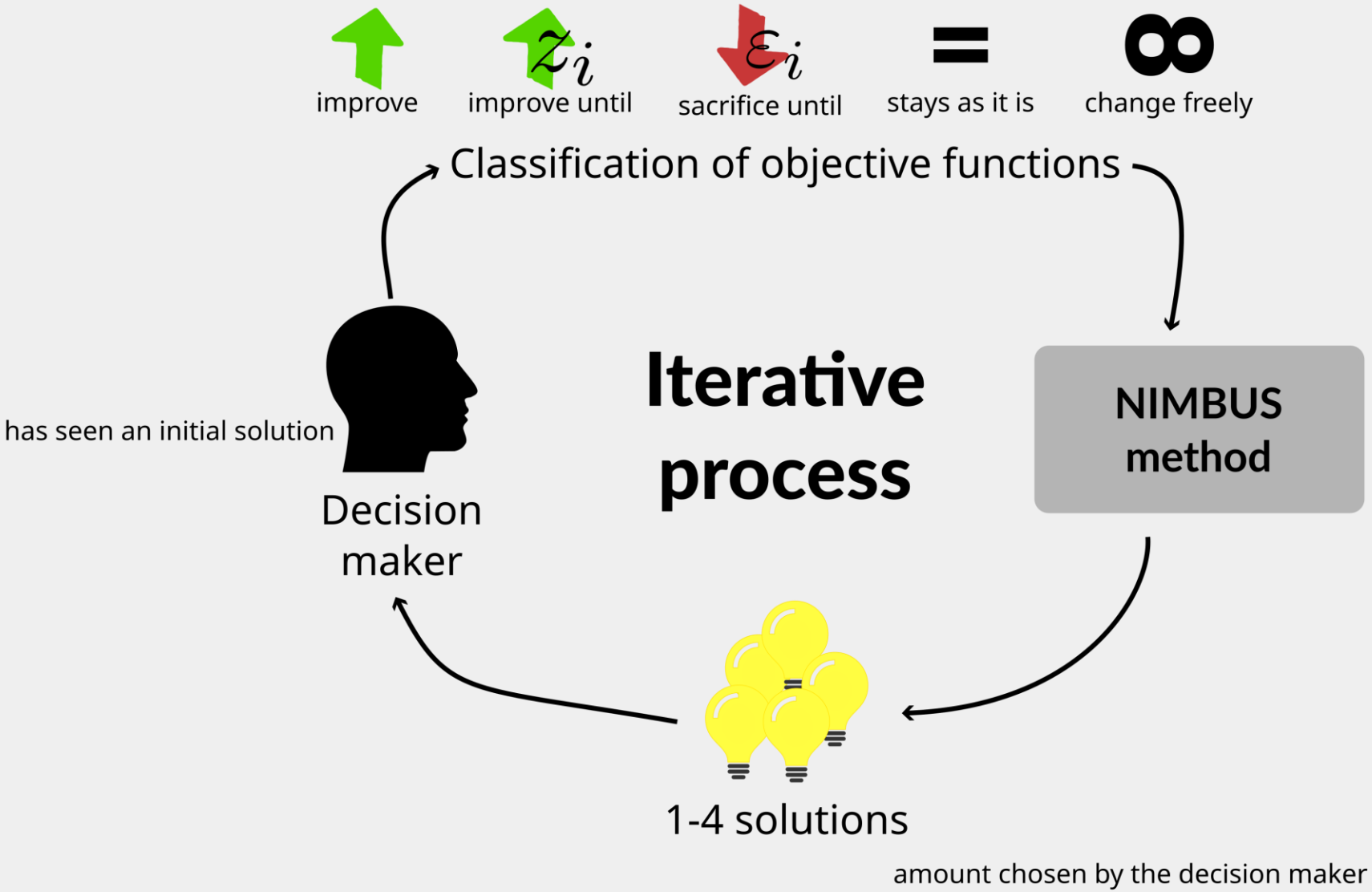
Miettinen, K. 1999. Nonlinear multiobjective optimization. Boston: Kluwer Academic Publishers.
Branke, J., Branke, J., Deb, K., Miettinen, K. & Slowiński, R. (Eds.) 2008. Multiob-ojective optimization: Interactive and evolutionary approaches. Springer.

The Reference Point method



Wierzbicki, A. P. (1980). The use of reference objectives in multiobjective optimization. In *Multiple criteria decision making theory and application: Proceedings of the third conference Hagen/Königswinter, West Germany, August 20–24, 1979* (pp. 468-486). Springer.
Wierzbicki, A. P. (1982). A mathematical basis for satisficing decision making. *Mathematical modelling*, *3*(5), 391-405.

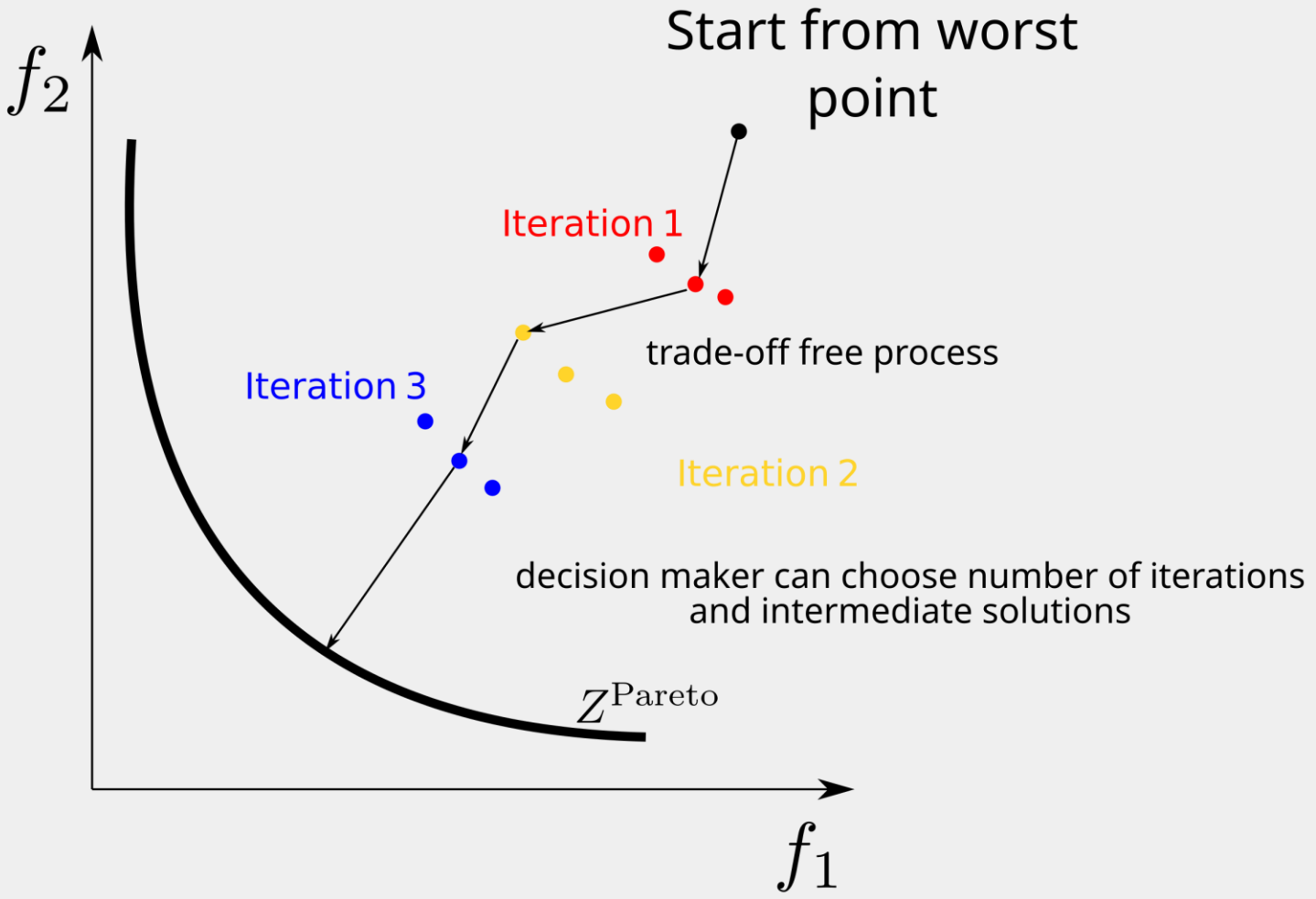
The (simplified) NIMBUS method



Miettinen, K., & Mäkelä, M. M. (2006). Synchronous approach in interactive multiobjective optimization. *European Journal of Operational Research*, *170*(3), 909-922.

The E-NAUTILUS method

Iterative process



Ruiz, A. B., Sindhya, K., Miettinen, K., Ruiz, F., & Luque, M. (2015). E-NAUTILUS: A decision support system for complex multiobjective optimization problems based on the NAUTILUS method. *European Journal of Operational Research*, *246*(1), 218-231.

Miettinen, K., & Ruiz, F. (2016). NAUTILUS framework: towards trade-off-free interaction in multiobjective optimization. Journal of Business Economics, 86, 5-21.

Comparing interactive methods

- Empirical research with decision makers is essential for accurately capturing human characteristics when comparing interactive methods.
- Recent studies lack such comparisons, and older studies have shortcomings, such as non-reproducibility because of missing experimental details (i.e., not reporting questions utilized) and reliance on one author simulating decision maker responses.
- This lack of detail prevents the replication of these experiments, as it is unclear how the studied phenomena were operationalized for measurement.

Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.
Afsar, B., Silvennoinen, J., Misitano, G., Ruiz, F., Ruiz, A. B., & Miettinen, K. (2023). Designing empirical experiments to compare interactive multiobjective optimization methods. Journal of the Operational Research Society, 74(11), 2327–2338.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Experiment design

- Questionnaire
- User interface
- Execution and practicalities

Questionnaire

- Based on desirable properties identified for interactive multiobjective optimization methods by Afsar et al. 2021.
- In the first study (Afsar et al. 2023), the questionnaire was designed and piloted in a small-scale experiment with two interactive methods considered (within-subjects).
- In the second study (Afsar et al. 2024), it was refined and applied in a much larger experiment with three interactive methods considered (between-subjects).

Research questions
RQ1: Cognitive load: How extensive is the cognitive load of the whole solution process?
RQ2: Capturing preferences and responsiveness: How well does the method capture and respond to the DM's preferences?
RQ3: Satisfaction and confidence: Is the DM satisfied with the overall solution process and confident with the final solution?

Afsar, B., Miettinen, K., & Ruiz, F. (2021). Assessing the Performance of Interactive Multiobjective Optimization Methods: A Survey. ACM Computing Surveys, 54(4), 1–27.
Afsar, B., Silvennoinen, J., Misitano, G., Ruiz, F., Ruiz, A. B., & Miettinen, K. (2023). Designing empirical experiments to compare interactive multiobjective optimization methods. *Journal of the Operational Research Society*, *74*(11), 2327-2338.
Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

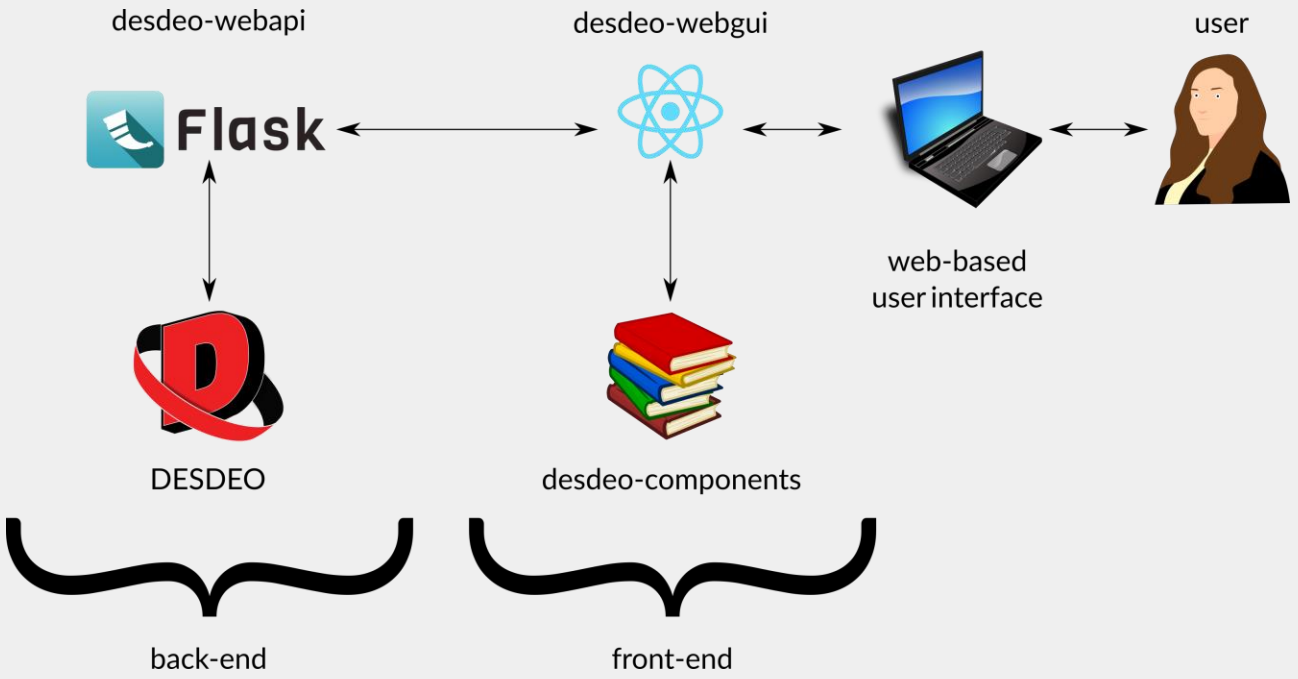
Questionnaire

- Each desirable property is assessed by one or more questions on a
  - Likert scale: "**It was easy to explore solutions with different conflicting values of the objective functions.**" (RQ2: "The method easily changes the area explored as a response to a change in the preference information given by the decision maker.")
  - 5-point semantic differential scale: "**What degree of conflict do you think exists between each pair of objectives? Among f1 and f2? Among f1 and f3? Among f2 and f3?**" (RQ3: "The method sets as low a cognitive burden on the decision maker as possible.)
  - Open ended: "**Did some solution(s) surprise you? Why?**" (RQ3: The method does not miss any Pareto optimal solution that is more preferred (with a given tolerance) for the decision maker than the one chosen.")
- Some questions were adapted from the NASA task load index (NATA-TLX), a validated measurement.
- Total of 29 questions. Some to be answered during an interactive method, some after the decision maker has found their final solution.

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50, 904–908.
Afsar, B., Miettinen, K., & Ruiz, F. (2021). Assessing the Performance of Interactive Multiobjective Optimization Methods: A Survey. ACM Computing Surveys, 54(4), 1–27.
Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

User interface

- Web-based interfaces for each method (The Reference Point method, NIMBUS, E-NAUTILUS).
- Accessible with any modern web-browser.
- Questionnaire integrated in the interface, no need to switch to other applications. Possibiilty to log user information.
- We utilized the DESDEO framework and its implementations of the interactive methods under study.



Misitano, G., Saini, B. S., Afsar, B., Shavazipour, B., & Miettinen, K. (2021). DESDEO: The Modular and Open Source Framework for Interactive Multiobjective Optimization. IEEE Access, 9, 148277–148295.
Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

User interface

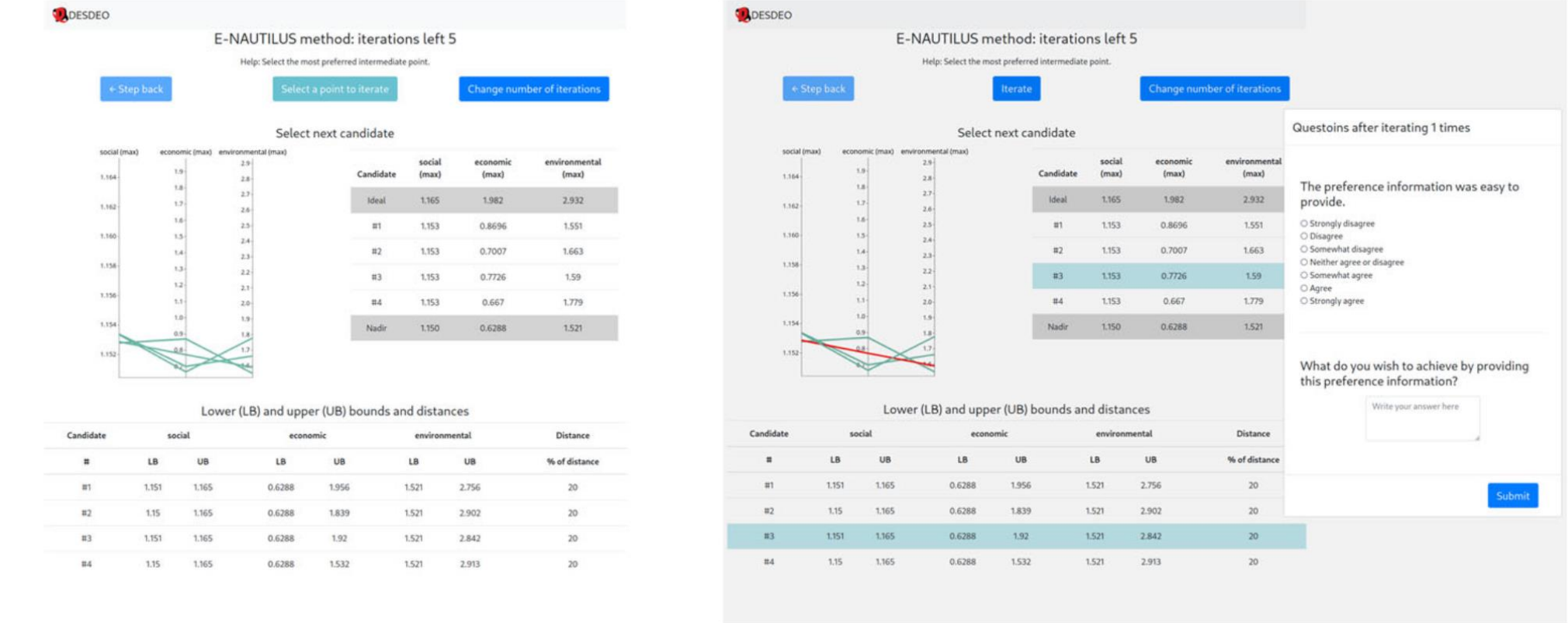
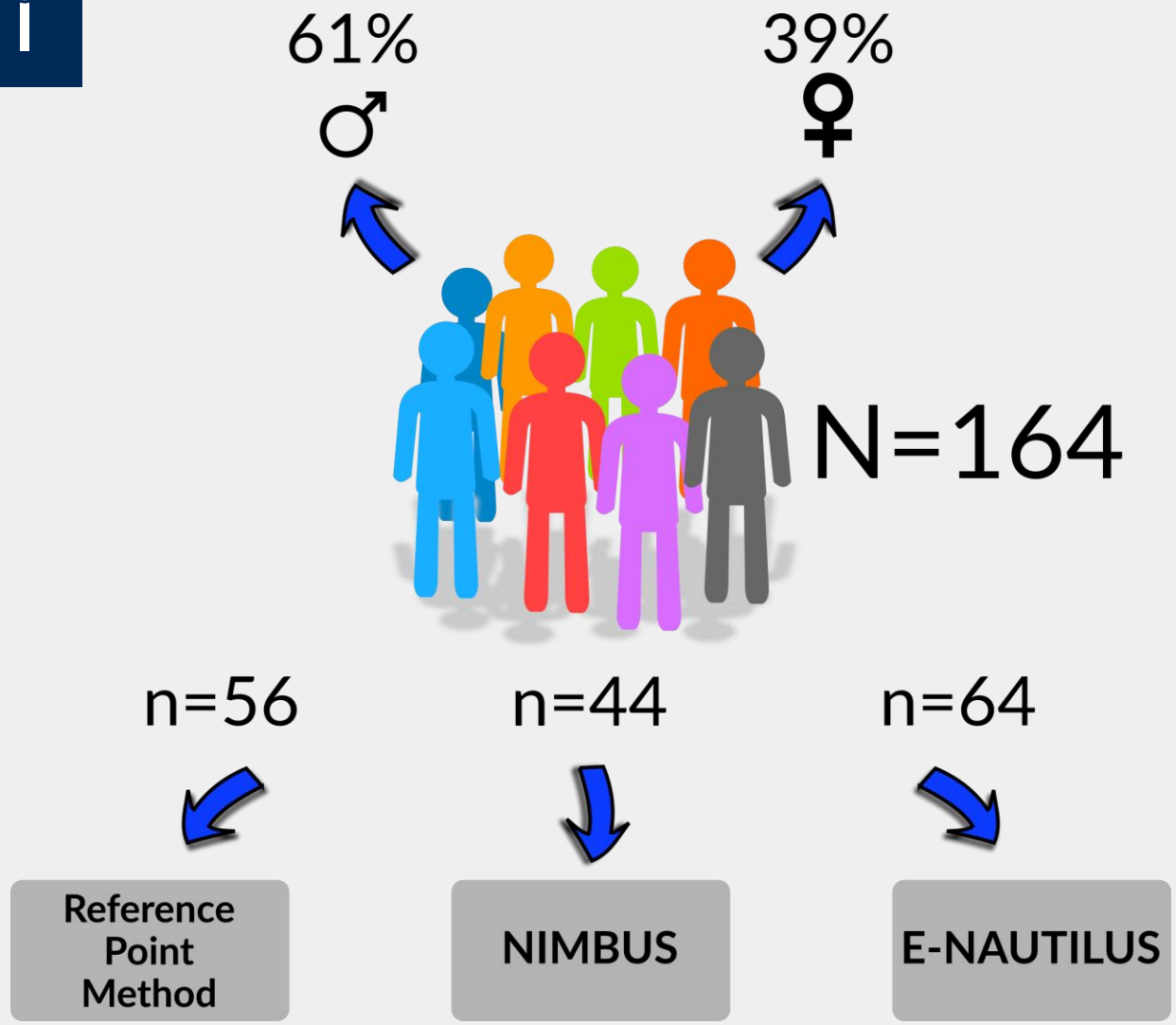


Fig. 1 Left: The UI of the E-NAUTILUS method. Right: Questionnaire items related to the given preferences as shown to the participant

Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

Execution and practicalities (in the second study)



- Between-subjects study
- Student participants from University of Malaga, Faculty of Economics and Business Studies.
- Age range 18-28, mean=19, and SD=2.2.
- Participants were briefed on the interactive methods and the problem being solved in advance to mitigate effects from being unfamiliar with the method or the problem.
- Particiapnts were also provided with suplementary material prior to the experiment to allow them to think carefully about their preferences before the study.

Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

Execution and practicalities

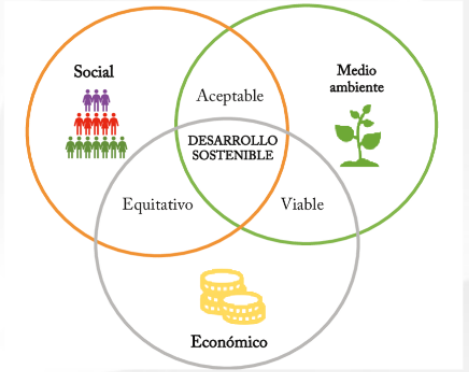
Problema de Programación Multiobjetivo para Analizar la Sostenibilidad de España

Indicadores individuales de la Agenda 2030 para el Desarrollo Sostenible de las Naciones Unidas

$\mathbf{x} = (x_1, \ldots, x_{11})^T$ es el vector de decisión con los 11 indicadores individuales más significativos

Tres dimensiones de la sostenibilidad

$f_j(\mathbf{x})$ = indicador compuesto que representa la dimensión j, siendo $j = 1$ (social), 2 (económica), y 3 (mediambiental)



Problema de Optimización Multiobjetivo

$$\begin{array}{ll} \text{Maximizar} & \{f_1(\mathbf{x}), f_2(\mathbf{x}), f_3(\mathbf{x})\} \\ \text{subject to} & \mathbf{x} \in S \end{array}$$

Significado de los valores de los indicadores compuestos para las tres dimensiones (f_1, f_2 y f_3)

Valores	Rendimiento global de la dimensión
[0, 1]	Entre el peor valor y el P_{25} de los países de la UE
(1, 2]	Entre el P_{25} y el P_{50} de los países de la UE
(2, 3]	Entre el P_{50} y el P_{75} de los países de la UE
(3, 4]	Entre el P_{75} y el mejor valor de los países de la UE

Situación de España en 2017

	Social (f_1)	Economico (f_2)	Medioambiental (f_3)
Ideal	1.1653	1.9822	2.9324
Actual	1.1550	1.3053	1.9281
Nadir	1.1504	0.6287	1.5214

- Decisor = Político intereseado en encontrar el mejor compromiso posible entre las tres dimensiones para mejorar la sostenibilidad en España.
- ¿Qué sacrificios son necesarios entre las tres dimensiones para alcanzar una situación mejor?
- ¿Cuáles serían los mejores valores que se podrían lograr para las tres dimensiones, en base a tus preferencias?

- Problem with three objective functions to be maximized.
- The three objective functions consists of a social one, an economical one, and an environmental one.
- The goal was to choose a problem that would matter to the students, so that they must make actual choices.
- Based on the sustainable development goals setup by the United Nations.

Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.
https://international-partnerships.ec.europa.eu/policies/sustainable-development-goals_en (accessed 12.6.2024)

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Results

- Analysis
- Answering RQ1
- Answering RQ2
- Answering RQ3
- Which method is best?

Analysis

- To analyse the results, both a quantitavive and qualitative approach was used.
  - For quantitative results (Likert scale and 5-point semantic differential scale) we adopted the Kruskal-Wallis test. We report the mean and standard deviation alongside *p*-values.
  - For qualitative results, we employed a qualitative content analysis approach.

Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

RQ1 - Cognitive load: How extensive is the cognitive load of the whole solution process?

Key Takeaways

- **E-NAUTILUS:** offers a balanced approach with lower effort, frustration, and iterations, although it requires more time.
- **NIMBUS:** preferred for satisfaction but involves higher frustration and effort.
- **The Reference Point method (RPM):** quick but less satisfactory and mentally demanding.

Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

RQ2 - Capturing preferences and responsiveness: How well does the method capture and respond to the DM's preferences?

Key Takeaways

- **E-NAUTILUS:** easiest for providing preferences, learning to use, and exploring conflicting solutions; better at returning to previous solutions.
- **NIMBUS:** best at generating solutions that reflect preferences and reacting to changes; good at capturing and responding to preferences.
- **RPM:** adequate performance but not exceptional in any specific area; good at representing user preferences but with some difficulties due to trade-offs.

Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

RQ3 - Satisfaction and confidence: Is the DM satisfied with the overall solution process and confident with the final solution?

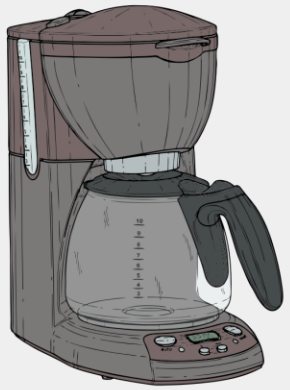
Key Takeaways

- **E-NAUTILUS:** highest overall satisfaction, easy to learn and return to solutions; participants found solutions similar to imagined.
- **NIMBUS:** best for final solution satisfaction and confidence; moderate knowledge gain and surprise levels.
- **RPM:** adequate performance with lower satisfaction and more surprises in final solutions; less knowledge gain.

Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

Which method is best?

E-NAUTILUS



Best For: Low mental effort, exploring trade-offs, clear understanding of overall solutions.
Recommended When: Ease of use and minimal frustration are priorities.

NIMBUS

Best For: Fine-tuning final solutions, accurate preference reflection.
Recommended When: High satisfaction with final solution and detailed control are needed.

RPM

(instant coffee)

Best For: Quick solutions with less complexity in preference information..
Recommended When: Time efficiency is crucial, but willing to accept lower satisfaction.

Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Conclusions

Conclusions

- Comparing interactive multiobjective optimization methods is challenging.
- We have designed, implemented, piloted, and applied our method.
- We have shared openly our method to others to use in their own experiments.
- More experiments are needed to truly say anyting definitieve on which interacitve method is suitable for different situations. **More experiments are needed!**
- We are also exploring on how to compare interactive mehtods when the method is switched between iterations, and when multiple decision makers are involved in a group decision-making setting.

Afsar, B., Silvennoinen, J., Ruiz, F., Ruiz, A. B., Misitano, G., & Miettinen, K. (2024). An experimental design for comparing interactive methods based on their desirable properties. Annals of Operations Research.

To support your endeavors

- DESDEO has played an important supporting role in enabling the works discussed in this presentation.
- DESDEO is currently going through a complete overhaul, which will make it more usable and welcoming to wild new ideas, including explainability.

The Multiobjective Optimization Group

We regularly post about our activities on LinkedIn!

Misitano, G., Saini, B. S., Afsar, B., Shavazipour, B., & Miettinen, K. (2021). DESDEO: The modular and open source framework for interactive multiobjective optimization. *IEEE Access*, *9*, 148277-148295.

Annals of Operations Research
https://doi.org/10.1007/s10479-024-05941-6

ORIGINAL RESEARCH

An experimental design for comparing interactive methods based on their desirable properties

Bekir Afsar[1] · Johanna Silvennoinen[1] · Francisco Ruiz[2] · Ana B. Ruiz[2] · Giovanni Misitano[1] · Kaisa Miettinen[1]

Received: 17 April 2023 / Accepted: 9 March 2024

Abstract

In multiobjective optimization problems, Pareto optimal solutions representing different tradeoffs cannot be ordered without incorporating preference information of a decision maker (DM). In interactive methods, the DM takes an active part in the solution process and provides preference information iteratively. Between iterations, the DM can learn how achievable the preferences are, learn about the tradeoffs, and adjust the preferences. Different interactive methods have been proposed in the literature, but the question of how to select the best-suited method for a problem to be solved remains partly open. We propose an experimental design for evaluating interactive methods according to several desirable properties related to the cognitive load experienced by the DM, the method's ability to capture preferences and its responsiveness to changes in the preferences, the DM's satisfaction in the overall solution process, and their confidence in the final solution. In the questionnaire designed, we connect each questionnaire item to be asked with a relevant research question characterizing these desirable properties of interactive methods. We also conduct a between-subjects experiment to compare three interactive methods and report interesting findings. In particular, we find out that trade-off-free methods may be more suitable for exploring the whole set of Pareto

Link to paper (open access!)

The Multiobjective Optimization Group

Research Council of Finland

Grant number 322221