Advances in Reliability, Safety and Security, Part 7 – Kołowrocki, Magryta-Mut (eds.)
© 2024 Polish Safety and Reliability Association, Gdynia, ISBN 978-83-68136-19-7 (printed), ISBN 978-83-68136-06-7 (electronic)

Advances in Reliability, Safety and Security

ESREL 2024
Monograph Book Series

Effect Of Weights And Criteria Uncertainty On Scoring In Multi Criteria Decision Analysis Ranking Problem: Application To Electricity Supply Resilience Index

Matteo Spada[a], Peter Burgherr[b], Flaminia Catalli[c], Eleftherios Siskos[d], Patrick Gasser[e]

[a]*Zurich University of Applied Sciences (ZHAW), Winterthur, Switzerland*
[b]*Paul Scherrer Institute (PSI), Villigen PSI, Switzerland*
[c]*wetransform GmbH, Darmstadt, Germany*
[d] *School of Production Engineering and Management, Technical University of Crete, Chania, Greece*
[e]*AFRY Management Consulting AG, Zurich, Switzerland*

Abstract

The resilience of an energy system and especially the national electricity supply is a complex and multidimensional concept, which is receiving growing attention in Europe. The increasing risks of extended electricity supply disruptions and/or severe electricity price fluctuations are stressing the need for an assessment of the European countries' electricity supply resilience. This study presents an extension of the Electricity Supply Resilience Index (ESRI), which was developed to benchmark countries' electricity supply resilience. The extension regards the inclusion of robustness analysis into the already performed sensitivity analysis in the ESRI framework. In fact, ESRI is built considering the sensitivity of the scoring/ranking of alternatives to the combination of normalization/aggregation functions, used to assess the performance of the countries. In this study, the extension is made in two ways. First, the scoring/ranking of the alternatives is analyzed to assess the effects of the criteria weights on the ESRI. Second, the scoring/ranking of the alternatives is analyzed for the case of uncertainty in the criteria performance the achieve. In total, 35 European countries are evaluated and ranked according to their performance on 17 evaluation criteria. Results show that the weighting has minimal effects on the final scoring/ranking of alternatives, while the presence of criteria uncertainty could have a significant effect on the final scoring/ranking.

Keywords: resilience, ENTSO-E, benchmark, ESRI, multi-criteria decision analysis (MCDA), stochastic analysis

1. Introduction

In our modern society, electricity is one of the most important commodities to foster the economic development and wealth of a country. Governments are increasingly aware of the need to improve the energy efficiency of electricity production, as it leads to better supply security and reduced greenhouse gas emissions. However, although the global average efficiency is expected to continuously improve with the correct investments (IEA, 2023), major electricity supply disruptions still happen causing great damage to the economy of a country (Kröger, 2017). As shown in the past decades, these major disasters shows that not all hazard and threats can be averted (Cimellaro, 2018), which calls for the analysis of the resilience of critical infrastructures. Indeed, resilience aims at minimizing the impact of adverse consequences by defining pre- and post-event strategies, making outages less likely or smaller in extent (Gasser et al. 2019).

Based on these premises, a resilient electricity supply is fundamental to guarantee a well-functioning modern society. In this regard, one of the key interests of policymakers is to assess how their country performs compared to others. Hence, resilience benchmarks are used to assess the progress of a country over a period and compare its performance against its peers (Gasser et al. 2020). Such focused assessments can have significant practical impact and drive socio-economic and political progress, influencing and driving the development of related initiatives, policies, and technological research (Siskos et al. 2014). The multi-dimensionality of electricity supply resilience, and the conflicting viewpoints and interests of the different stakeholders require an accurate and transparent benchmarking framework. In this context, Multi-Criteria Decision Analysis (MCDA) offers the appropriate theoretical background and principles, as well as a multitude of tools and techniques to analyze such a complex problem.

In the past, MCDA was applied in the context of countries' benchmarking by using the so-called composite indicators (CI), which combine the individual criteria into a score/rank. For this purpose, different operational research methods are used to build such indices, for example the analytical hierarchy process (AHP) (Wu et al., 2012), weighted averages (Valdés, 2018) and multi-attribute value/utility theory (MAVT/MAUT) (Pohekar and Ramachandran, 2004) to name some. Recently, such a CI was developed in the context of the electricity supply resilience within the Future Resilient Systems (FRS) program at the Singapore-ETH Centre (SEC) to assess the resilience of electricity supply in countries by considering different dimensions of resilience (Gasser et al., 2020). The Electricity Supply Resilience Index (ESRI) was created not just for evaluating the resilience of countries' electricity supply, but also for examining how the sensitivity of country scores to normalization and aggregation methods impacts the analysis. This sensitivity, as discussed by Gasser et al. (2020), can lead to variations in rankings to some degree (Narula and Reddy, 2015). Subsequently, the methodological framework of ESRI has been expanded in two ways. First, to explore the impact of criteria correlation on stakeholder-assigned weights and to propose an optimization approach to mitigate this influence, as discussed by Lindén et al. (2021). Second, to investigate the influence of interacting criteria and offer a reliable elicitation method, as discussed by Siskos and Burgherr (2022).

In the past, it has been demonstrated that both scores and rankings are susceptible not only to the normalization and aggregation techniques used for evaluation but also to the weight preferences assigned to various criteria in constructing the Composite Indicator (CI), as discussed by Dobbie and Dail (2013). Additionally, criteria uncertainty could exert an influence on the ultimate scoring and ranking of alternatives, as highlighted by Groothuis-Oudshoorn et al. (2017).

Based on these premises, starting from the ESRI methodological framework that focused on analyzing the sensitivity of the scoring/ranking of countries to the combination normalization/aggregation functions used in the construction of the CI (Gasser et al., 2020), in this study two additional robustness analyses related to the (i) use of different weighting profiles, and (ii) the presence of uncertainty in criteria are considered. In this context, the following research questions are posed:

- What are the effects on the final scoring/ranking of alternatives when different weighting profiles are considered?
- What are the effects on the final scoring/ranking of alternatives when uncertainty in the criteria scoring is present in the initial dataset?

The remainder of the paper is structured as follow. In section 2, the ESRI index and the proposed extension, its dimensions, and criteria are presented. In section 3 the method to assess the sensitivity and robustness of the CI is described, while in section 4 the results of the study are shown. Finally, in section 5 conclusive remarks are given.

2. The Electricity Supply Resilience Index (ESRI)

The Electricity Supply Resilience Index (ESRI) is a CI that is a mathematical combination of individual criteria that together act as a proxy for the phenomena being measured (Mazziotta and Pareto, 2013). By combining multiple variables - using MCDA techniques - CIs can quantitatively assess and rank the performance of alternatives across multidimensional concepts, providing flexible tools to support decision making when more than one criterion is considered (Cinelli et al. 2014). Therefore, the ESRI is developed according to the widely accepted ten-step methodological framework for good practice in the construction of CIs (Organisation for Economic Co-operation and Development (OECD), 2008):

1. Theoretical framework
2. Data selection
3. Imputation of missing data
4. Multivariate analysis
5. Normalisation
6. Weighting and aggregation
7. Uncertainty and sensitivity analysis
8. Back to the data
9. Links to other criteria

10. Visualisation of results

The first version of ESRI, applied to 140 countries worldwide (Gasser et al., 2020), considered a set of 12 criteria that describe the multifaceted nature of the four dimensions of resilience, according to Heinimann and Hatfield (2017):

- Resist: represents the ability of the system to withstand disturbances within acceptable levels of degradation.
- Restabilise: illustrates the ability to limit degradation of performance and restore key functions.
- Rebuild: describes the process of restoring system performance to normal.
- Reconfigure: characterises the changes to the biophysical architecture/topology of the system to make it more fault tolerant.

The ESRI was expanded by augmenting the number of criteria from 12 to 17 to provide a more comprehensive depiction of resilience's multifaceted nature. Additionally, the number of resilience dimensions was reduced from four to three to ensure an equitable representation of the dimensions, while preventing the simultaneous assignment of a criterion to multiple resilience dimensions (Siskos and Burgherr, 2022). In addition, the number of countries was reduced from 140 to 35 due to the lack of worldwide information for some of the additional criteria considered in the ESRI extension. In this context, the resilience of electricity supply was assessed for European countries, and in particular for the 35 countries that are members of the ENTSO-E group, namely Albania, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Slovenia, Slovakia, Spain, Sweden, Switzerland and the United Kingdom (Siskos and Burgherr, 2022).

2.1. ESRI extension

This study focuses on the last version of the ESRI index methodology, and how it can be extended with regard to steps 6 and 7 of the ten-step methodological framework for good practice in the construction of CIs (Organisation for Economic Co-operation and Development (OECD), 2008). Particular attention is given to the effect of both the different weighting schemes, and the uncertainty arising from the criteria on the scoring/ranking of alternatives.

In this section the resilience dimensions and criteria used in this study are summarised in sections 2.1.1 and 2.1.2. Furthermore, section 2.1.3 presents the steps on how the input matrix for the MCDA analysis is developed, by first describing how the probabilistic distributions for the uncertainty criteria are established (section 2.1.3.1), followed by a multivariate analysis on the input matrix for the MCDA (step 4 in section 2), which assesses how consistent are the selected criteria to build an index (section 2.1.3.2). Finally, the input matrix for the MCDA analysis is presented (section 2.1.3.3).

2.1.1. Resilience Dimensions and Criteria

In this study, ESRI defined by three dimensions (Resist, Restabilize and Recover), and described by a total of 17 criteria (Siskos and Burgherr, 2022). The dimension Recover combines the Rebuild and Reconfigure dimensions from the initial ESRI version. This approach addresses the interrelated phases of system recovery and helps overcome challenges in specifying criteria separately for each phase (Siskos and Burgherr, 2022). The 17 criteria were subsequentially allocated to the three resilience dimensions to guide the decision-making process in a controllable, complete, measurable, non-redundant and concise manner. Specifically, an equal representation of the three resilience dimensions was achieved through similar numbers of criteria per dimension (Siskos and Burgherr, 2022). In the end, the three dimensions contain six, six and five criteria respectively (see Figure 1).

2.1.2. Data Preparation and MCDA Input Matrix

To build up the MCDA input matrix, different sources are used to collect the data used to assess each criterion (Table 1). The collected data are the raw information used in this study and are analysed and further processed to (i) assess the probabilistic distribution of the criteria under uncertainty (section 2.3.1), (ii) assess the coherence of the criteria set structure (section 2.3.2) to follow step 4 in section 2, (iii) build the input matrix to be used in the MCDA analysis (section 2.3.3).

2.1.2.1. Estimation of Criteria Uncertainty

In this section the analysis of each criterion in Figure 1 and the estimation of the probabilistic distributions describing a criterion uncertainty are presented. For a detailed description of each criterion used in this study, please refer to Siskos and Burgherr (2022).

As shown in Table 1, data to assess criteria C_2, C_6, C_{10}, C_{17} are collected for a specific year, either due to lack of data (C_6, C_{10}, C_{17}) or since there is no significant yearly variation of the criterion (C_2). In this context, the criteria volatility within a specific period cannot be analysed. Therefore, they are directly included in the input matrix of the MCDA without any further analysis.

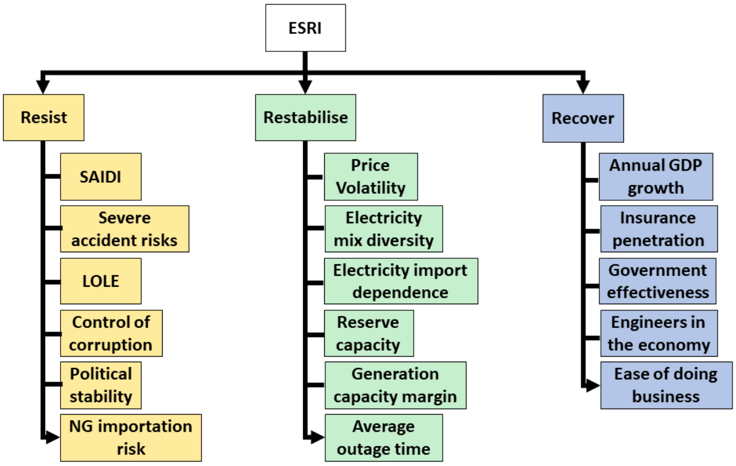


Fig. 1. The evaluation system of Electricity Supply Resilience Index (ESRI), modified from Siskos and Burgherr (2022).

On the other hand, criterion C_7 describes the volatility of electricity prices, which is why data for the years 2015-2021 is used to directly calculate the volatility of the electricity prices as input for the MCDA analysis, rather than to define a probabilistic distribution describing the uncertainty of this criterion.

Data to assess criteria C_1, C_4, C_5, C_8, C_9, C_{11}, C_{12}, C_{13}, C_{14}, C_{15}, C_{16} are collected for the period 2015-2021, while C_3 data are taken from the 2025, 2027 and 2030 reference scenarios published in ENTSO-E (2023). For each of these criteria and for all the alternatives considered (ENTSO-E countries), a visual analysis was first carried out to understand what kind of distribution each criteria could have followed and then, if necessary, an Akaike Information Criterion (AIC) was applied to each potential probabilistic distribution, e.g., normal, lognormal, Poisson, etc., to evaluate the model that best fits the data (Bozdogan, 1987). However, in this study, the visual analysis already indicated that a normal distribution was a good fit for all criteria uncertainty, so no further analysis was carried out. Based on these premises, the mean and standard deviation are estimated for each ENTSO-E country and for each of the abovementioned criteria. These represent the parameters in the input matrix for the MCDA analysis (see Table 1).

2.1.2.2. Coherence Criteria set structure

The fourth step in the framework for building a coherent CI is a multivariate analysis (section 2). The latter is carried out to assess the reliability of a set of criteria and their internal consistency to develop an index (Organisation for Economic Co-operation and Development (OECD), 2008). In this study, this quality of the dataset is measured using Cronbach's Alpha, which is the most widely used index to assess the reliability of a scale (Streiner, 2003). It measures how closely related a set of criteria are as a group. Cronbach's Alpha values below 0.7 indicate questionable internal consistency, while values above 0.9 indicate excessive redundancy between criteria. The present case study dataset has a Cronbach's Alpha of 0.74, which is within the desirable range for consistent composite scales developed for research purposes (0.7 to 0.9) (Streiner, 2003).

2.1.2.3. Multi-Criteria Decision Analysis (MCDA) Input Dataset

Based on the data collection process described in sections 2.1.2.1 and 2.1.2.2, the resulting input matrix for the MCDA analysis is shown in Table 1.

Table 1 indicates for each criterion whether it is described by a deterministic value ("Exact") or by a probabilistic distribution, which is "Normal" in all cases considered in this study, as described in section 2.1.2.1. The arrow indicates the polarity of a criterion. The polarity is important in the context of MCDA, as it defines whether a criteria should be minimised (↓), meaning that the criterion performs better the lower its value is, or maximised (↑), meaning that a higher value denotes a better performance.

Table 1. Input Matrix for the MCDA Analysis.

Polarity	↓		↓	↓		↑		↑		↓	↓	↑	
Distribution	Normal		Exact	Normal		Normal		Normal		Exact	Exact	Normal	
Collection Period	2015-2021		2021	2025, 2027, 2030		2015-2021		2015-2021			2015-2021	2015-2021	
Data Source	World Bank		Spada and Burgherr (2023), IEA	ENTSO-E (2023)		World Bank		World Bank		Siskos and Burgherr (2022)	Eurostat	IEA	
	C_1		C_2	C_3		C_4		C_5		C_6	C_7	C_8	
ENTSO-E Countries	System Average Interruption Duration Index (SAIDI) [h/year]		Severe accident risks [Fatalities/GWeyr]	Loss of Load Expectation (LOLE) [h/year]		Control of corruption [Index]		Political stability and absence of violence/ terrorism [%]		Risk of exporting countries [Index]	Volatility of electricity prices (households) [%]	Electricity mix diversity	
Albania	65.97	23.34	0.95	0.30	0.22	-0.57	0.11	0.21	0.17	0.00	0.13	0.00	0.01
Austria	0.67	0.03	0.01	1.10	0.37	1.45	0.10	1.02	0.22	0.28	0.03	0.41	0.04
Belgium	0.50	0.11	0.01	7.13	2.45	1.49	0.06	0.60	0.18	0.17	0.08	0.55	0.03
Bosnia and Herzegov	2.31	0.12	0.70	2.40	1.10	-0.50	0.16	-0.39	0.12	0.41	0.02	0.29	0.02
Bulgaria	5.76	0.55	0.02	0.67	0.45	-0.26	0.06	0.28	0.17	0.40	0.07	0.58	0.01
Croatia	4.34	0.74	0.01	0.00	0.00	0.11	0.07	0.65	0.06	0.22	0.08	0.40	0.07
Cyprus	0.45	0.11	0.05	0.00	0.00	0.76	0.31	0.51	0.09	0.00	0.19	0.20	0.05
Czech Republic	0.50	0.03	0.02	0.17	0.24	0.50	0.11	0.97	0.07	0.40	0.08	0.50	0.03
Denmark	0.51	0.01	0.00	18.30	4.68	2.26	0.10	0.91	0.04	0.15	0.03	0.35	0.02
Estonia	0.51	0.26	0.00	7.40	2.16	1.35	0.17	0.68	0.06	0.41	0.09	0.47	0.05
Finland	0.26	0.22	0.01	2.40	0.80	2.19	0.04	1.06	0.19	0.40	0.05	0.53	0.02
France	0.25	0.07	0.01	8.20	1.87	1.28	0.08	0.25	0.18	0.26	0.03	0.41	0.04
Germany	0.20	0.04	0.02	14.87	4.12	1.82	0.04	0.70	0.13	0.30	0.04	0.76	0.01
Greece	1.89	0.26	0.02	0.67	0.54	-0.05	0.10	-0.03	0.15	0.42	0.06	0.68	0.03
Hungary	2.77	0.15	0.01	4.17	1.64	0.10	0.11	0.74	0.07	0.40	0.07	0.54	0.03
Iceland	0.56	0.05	0.01	0.00	0.00	1.79	0.11	1.34	0.10	0.00	0.11	0.26	0.01
Ireland	0.45	0.25	0.02	9.43	10.52	1.59	0.08	0.93	0.07	0.17	0.07	0.48	0.01
Italy	0.76	0.38	0.02	11.27	1.36	0.23	0.19	0.41	0.08	0.39	0.07	0.56	0.02
Latvia	1.03	0.06	0.02	2.77	3.70	0.47	0.16	0.48	0.08	0.41	0.10	0.34	0.01
Lithuania	0.42	0.04	0.01	5.33	1.09	0.60	0.14	0.79	0.08	0.28	0.09	0.59	0.02
Luxembourg	0.33	0.04	0.01	14.87	4.12	2.02	0.09	1.30	0.10	0.22	0.06	0.48	0.05
Montenegro	32.53	5.61	0.79	0.00	0.00	-0.09	0.07	0.14	0.22	0.00	0.03	0.30	0.00
Netherlands	0.63	0.11	0.02	7.13	5.05	1.93	0.09	0.93	0.14	0.21	0.11	0.40	0.04
North Macedonia	6.42	0.78	0.54	1.43	0.47	-0.28	0.16	-0.13	0.24	0.41	0.03	0.46	0.05
Norway	1.00	0.36	0.03	0.00	0.00	2.15	0.08	1.16	0.12	0.00	0.12	0.09	0.02
Poland	1.22	0.07	0.03	0.73	0.90	0.64	0.07	0.66	0.21	0.33	0.09	0.49	0.07
Portugal	0.58	0.04	0.01	0.00	0.00	0.83	0.08	0.93	0.12	0.43	0.07	0.58	0.01
Romania	2.56	0.42	0.02	0.67	0.73	-0.18	0.11	0.27	0.21	0.38	0.09	0.71	0.01
Serbia	4.07	0.20	0.66	2.30	0.78	-0.39	0.07	-0.01	0.15	0.41	0.06	0.31	0.03
Slovak Republic	0.88	0.09	0.01	0.00	0.00	0.17	0.10	0.80	0.21	0.41	0.06	0.51	0.01
Slovenia	0.39	0.19	0.02	1.07	0.57	0.78	0.06	0.86	0.10	0.23	0.05	0.55	0.01
Spain	0.56	0.09	0.01	3.37	2.36	0.72	0.17	0.26	0.15	0.40	0.07	0.70	0.01
Sweden	0.58	0.02	0.01	6.07	1.58	2.13	0.06	1.01	0.08	0.18	0.07	0.47	0.02
Switzerland	0.21	0.01	0.01	0.33	0.34	2.03	0.08	1.29	0.10	0.19	0.04	0.41	0.01
United Kingdom	0.29	0.03	0.01	5.13	0.66	1.74	0.08	0.42	0.12	0.17	0.09	0.58	0.02

Polarity	↓		↑	↑		↓		↑		↑		↑		↑		↑
Distribution	Normal		Exact	Normal		Normal		Normal		Normal		Normal		Normal		Exact
Collection Period	2015-2021		2020	2015-2021		2015-2021		2015-2021		2015-2021		2015-2021		2015-2021		2019
Data Source	IEA		ENTSO-E (2017)	ENTSO-E IEA		World Bank		World Bank		Swiss Re Institute		World Bank		Eurostat		World Bank
	C_9		C_{10}	C_{11}		C_{12}		C_{13}		C_{14}		C_{15}		C_{16}		C_{17}
ENTSO-E Countries	Electricity import dependence		Reserves capacity [% of total peak load]	Generation Capacity Margin [%]		Average outage time [h]		Annual GDP growth [%]		Insurance penetration [%]		Government effectiveness [%]		Engineers in the economy [%]		Ease of doing business [Distance to Frontier]
Albania	1.07	0.26	0.06	0.33	0.12	1.98	0.34	2.74	2.84	0.50	0.01	-0.05	0.13	0.00	0.00	67.70
Austria	1.06	0.05	0.04	0.25	0.08	1.04	0.08	1.24	2.85	0.29	0.03	1.51	0.05	0.01	0.00	78.70
Belgium	1.11	0.16	0.04	0.21	0.04	1.08	0.04	1.45	2.62	0.46	0.08	1.34	0.22	0.01	0.00	75.00
Bosnia and Herzegov	0.65	0.04	0.12	0.30	0.09	4.05	0.57	2.59	2.62	0.50	0.01	-0.69	0.25	0.01	0.00	65.40
Bulgaria	0.71	0.05	0.04	0.26	0.02	1.30	0.11	2.20	2.80	0.50	0.01	-0.03	0.14	0.00	0.00	72.00
Croatia	1.52	0.15	0.06	0.32	0.04	2.74	0.48	2.13	5.14	0.50	0.01	0.54	0.09	0.01	0.00	73.60
Cyprus	1.24	0.15	0.21	0.62	0.03	2.94	0.56	2.09	4.82	0.50	0.01	0.98	0.22	0.00	0.00	73.40
Czech Republic	0.81	0.02	0.11	0.32	0.02	1.62	0.20	1.94	2.94	0.50	0.01	0.99	0.06	0.01	0.00	76.30
Denmark	1.68	0.10	0.14	0.52	0.04	0.97	0.13	1.94	1.77	0.91	0.08	1.87	0.08	0.01	0.00	85.30
Estonia	7.60	0.94	0.14	0.30	0.12	1.99	0.44	2.93	2.50	0.50	0.01	1.12	0.14	0.00	0.00	80.60
Finland	1.69	0.11	0.10	0.07	0.01	0.83	0.69	0.82	1.81	0.92	0.05	1.97	0.13	0.01	0.00	80.20
France	0.81	0.02	0.02	0.25	0.03	1.12	0.23	1.06	3.26	0.77	0.04	1.34	0.10	0.01	0.00	76.80
Germany	1.09	0.05	0.06	0.37	0.02	0.92	0.16	1.11	1.70	0.48	0.02	1.50	0.14	0.00	0.00	79.70
Greece	1.04	0.04	0.08	0.38	0.02	1.28	0.22	0.02	4.78	0.01	0.01	0.30	0.10	0.00	0.00	68.40
Hungary	1.33	0.04	0.08	0.24	0.01	2.16	0.04	2.95	3.16	0.02	0.01	0.54	0.07	0.00	0.00	73.40
Iceland	0.94	0.00	0.10	0.00	0.01	1.07	0.33	2.95	3.64	0.50	0.01	1.47	0.06	0.00	0.00	79.00
Ireland	1.03	0.08	0.03	0.43	0.01	1.60	0.09	8.26	6.56	0.52	0.11	1.44	0.11	0.00	0.00	79.60
Italy	1.30	0.08	0.12	0.44	0.01	0.35	0.11	0.18	3.83	0.73	0.08	0.45	0.06	0.01	0.00	72.90
Latvia	1.26	0.15	0.07	0.26	0.03	1.68	0.46	2.81	2.13	0.50	0.01	0.91	0.11	0.00	0.00	80.30
Lithuania	3.00	0.56	0.21	0.59	0.06	1.21	0.36	3.29	1.54	0.50	0.01	0.98	0.09	0.00	0.00	81.60
Luxembourg	3.02	0.51	0.04	0.65	0.01	1.21	0.18	2.31	1.61	0.50	0.01	1.69	0.05	0.00	0.00	69.60
Montenegro	0.93	0.13	0.16	0.23	0.09	1.25	0.22	2.42	6.64	0.50	0.01	0.10	0.11	0.02	0.01	73.80
Netherlands	1.41	0.23	0.04	0.40	0.06	1.40	0.31	1.55	2.37	0.86	0.09	1.79	0.07	0.01	0.00	76.10
North Macedonia	1.17	0.06	0.12	0.41	0.02	0.53	0.09	2.02	2.97	0.50	0.01	-0.06	0.08	0.01	0.00	80.70
Norway	0.84	0.04	0.04	0.23	0.04	0.85	0.14	1.74	1.34	0.28	0.02	1.86	0.05	0.01	0.00	82.60
Poland	1.87	0.06	0.04	0.54	0.03	1.07	0.11	3.53	2.43	0.50	0.01	0.54	0.18	0.00	0.00	76.40
Portugal	1.13	0.11	0.03	0.30	0.11	0.82	0.06	1.14	4.07	0.45	0.11	1.09	0.11	0.00	0.00	76.50
Romania	0.79	0.06	0.13	0.25	0.01	0.96	0.07	3.39	3.03	0.50	0.01	-0.06	0.15	0.01	0.00	73.30
Serbia	0.78	0.02	0.07	0.11	0.03	1.16	0.12	2.33	2.57	0.50	0.01	-0.02	0.08	0.01	0.00	75.70
Slovak Republic	1.05	0.07	0.05	0.20	0.02	1.38	0.16	2.22	2.22	0.50	0.01	0.64	0.14	0.01	0.00	75.60
Slovenia	0.85	0.05	0.08	0.12	0.02	1.54	0.67	2.41	3.53	0.50	0.01	1.06	0.06	0.01	0.00	76.50
Spain	1.06	0.08	0.03	0.40	0.03	0.70	0.06	0.98	4.60	0.35	0.02	1.03	0.11	0.00	0.00	77.90
Sweden	0.85	0.03	0.12	0.21	0.06	1.26	0.17	2.02	1.99	0.55	0.05	1.69	0.09	0.00	0.00	82.00
Switzerland	0.86	0.06	0.09	0.25	0.06	1.22	0.16	1.66	1.53	0.66	0.08	1.97	0.08	0.01	0.00	76.60
United Kingdom	1.17	0.13	0.07	0.36	0.07	1.99	0.19	1.59	4.33	0.94	0.07	1.46	0.15	0.01	0.00	83.50

3. Method

The construction of the ESRI for each of the considered ENTSO-E countries is based on the combination of (i) a normalisation method, which allows to assign all criteria on a common scale, so that they can be compared to each other, (ii) a weighting profile, which defines how much importance is given to each individual criterion, and (iii) an aggregation method, which combines the normalised criteria data with their respective weights into an index to assess the score of each alternative (steps 5 and 6 in section 2).

In this study, the sensitivity of the scores to a set of normalisation/aggregation combinations to construct the ESRI index is considered (section 3.1), following the approach of Gasser et al. (2020). First, all criteria in the input matrix were considered to be deterministic, i.e., only the means of the uncertainty criteria are taken from the input matrix, and the ESRI scores for equal weights are compared to weighting profiles sampled from a uniform distribution (section 3.2). Second, the results for an equal weighting profile are compared for the two cases that all the criteria are deterministic against the combination of deterministic and criteria uncertainty as shown in Table 1 (section 3.2).

To compute and assess these comparisons, this study makes use of a novel Python module, ProMCDA, developed by Catalli and Spada (2023), which allows to combine different normalisation and aggregation functions, sample weights, and take into account criteria uncertainty.

3.1. Normalization and aggregation methods

Many normalisation methods and aggregation functions are reported in the literature (Organisation for Economic Co-operation and Development (OECD), 2008). Therefore, the scoring and ranking of the alternatives under interest depends on which combination of normalisation and aggregation is used (Gasser et al., 2020). To assess the sensitivity of each country's ESRI, multiple rankings are derived by combining four normalisation methods (rank, standardisation, min-max, target) and four aggregation functions (additive, geometric, harmonic, minimum). The normalization methods reflect various preferences of decision-makers, ranging from solely acknowledging the ordinal nature of the data to considering quantitative disparities between performances, known as the cardinal character. Furthermore, the aggregation functions represent different levels of compensation between criteria, i.e., whether a decision maker is willing to allow a low performance of one criteria to be fully, partially, or not at all compensated by other criteria.

As shown in the literature, due to the nature of normalisation methods and aggregation functions, not all combinations are meaningful, and some combinations are redundant (Cinelli et al., 2020). In this study only 13 combinations are considered:

- Additive aggregation with Min-Max, Target, Standardized and Rank normalisation methods.
- Geometric mean aggregation with Min-Max, Target, Standardized and Rank normalisation methods.
- Harmonic mean aggregation with Min-Max, Target, Standardized and Rank normalisation methods.
- Minimum with Standardized normalisation method.

Therefore, in this study different criteria compensation are considered from the full compensatory aggregation method (Additive) to the least compensatory method (Minimum), where no compensation is allowed. Furthermore, is important to note that the Min-Max, Target, and Standardized normalisations for the Geometric and Harmonic aggregation methods are rescaled to avoid criteria with zero values, which are not allowed in the calculation of these means.

3.2. Robustness Analysis

The aim of this study is to understand the impact of the weighting schemes combined with the sensitivity described in section 3.1 on the final scores of the alternatives, as well as the impact of considering criteria uncertainty in the input matrix on the final scores of the alternatives.

To perform the robustness analysis, the ProMCDA Python module allows random sampling of either the weights or the criteria values using a Monte Carlo method. In ProMCDA, randomness is not allowed for either the weights or the criteria simultaneously, to make the results as transparent as possible. Randomness in the weights is applied by sampling a selected number of weighting profiles, corresponding to the number of Monte Carlo runs in ProMCDA, from a uniform distribution [0-1]. Additionally, since the weights should sum to 1, each generated weighting profile is checked and if the sum is not 1, the weights are normalised. On the other hand, the robustness analysis for the criteria is performed by assessing the probabilistic distribution that best describes the criteria. This means that if a criterion is characterised by a distribution described by, for example, 2 parameters as the Normal distribution, two columns should be allocated in the input matrix, as shown in Table 1 of this study. Once the probabilistic distribution for each criterion is selected and the input parameters are in place in the input matrix, ProMCDA randomly samples ***n*** values, where ***n*** is the number of Monte Carlo runs, of each criteria per alternative from the given distribution and assesses the score of the alternatives, considering the robustness at the criterion level.

4. Results

In this section, the comparisons of ESRI with equal weights and sample weights, and the one with deterministic and uncertainty criteria are presented. In this context, the base case scenario is the sensitivity analysis (considering the 13 combinations of normalisation/aggregation functions, see section 3) with equal weights for all criteria. A uniform weighting scheme is applied because it does not introduce further elements that could affect the ranking of the alternatives and does not include subjective preferences (weights). Equal weights also represent the most common profile for such a comparison (El Gibari et al. 2018). Furthermore, in this case the criteria are considered deterministic, i.e., without uncertainty (section 3).

The base case scenario is then compared with two additional cases. In the first case, labelled "sampled weights," all criteria are treated as deterministic. The same combinations of normalization and aggregation functions are employed, and weights are sampled from 10,000 weighting profiles. The selection of 10,000 profiles followed a convergence test, indicating robust results after this number of runs (see section 4.1). In the second case, labelled "uncertain criteria," the same normalization and aggregation combinations are used, with equal weights assigned to all criteria. However, criteria uncertainty is considered, with their values sampled 100,000 times from a normal distribution, as depicted in Table 1 (section 4.2). The choice of 100,000 samples followed a convergence test.

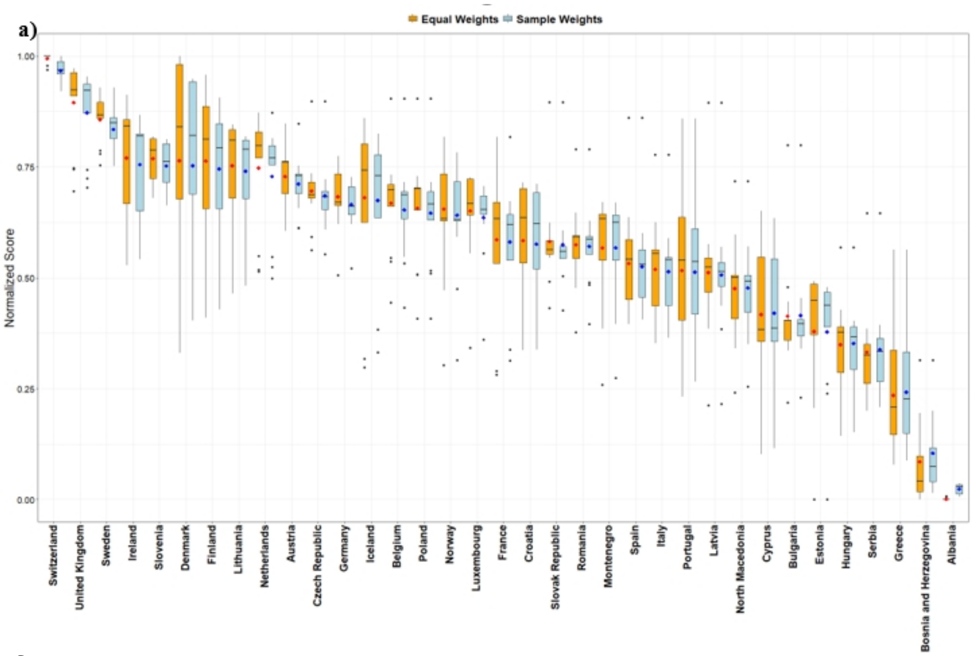
4.1. Equal Weights vs. Sample Weights

Figure 2a presents the comparison of the ESRI between the base case scenario and “sampled weights” case. The countries are classified in descending order, according to the ESRI mean score, estimated using the base case scenario. Distinct country groupings can be observed for the base case scenario, with few countries scoring at the top or at the bottom across all combinations of normalisation/aggregation methods. Switzerland and UK have top performing criteria (i.e., 5 and 4, respectively) and even their worst performing criteria still outperform those from many other countries. On the other hand, Bosnia & Herzegovina, and Albania do not have enough well-performing criteria to compensate the lower-performing ones; they thus score at the bottom, irrespectively of the combinations considered. Therefore, for the top- and bottom performing countries, different normalization methods and aggregation functions that allow different levels of compensation between the criteria do not have much influence on the final scores. However, the scores tend to be more variable for average-performing countries. In this context, the large disparity of the scores proves that the final rankings can have a strong dependency on the considered normalization/aggregation combination, as shown for Denmark, Finland, Lithuania, Portugal, Cyprus and Greece.

The ESRI case “sampled weights” shows similar results to the base case scenario. In particular, the distributions of the scores not considering the outliers are generally similar, except for some slight reductions (e.g., Denmark, Finland, Austria) or enlargements (e.g., Switzerland, Albania) of the size of the boxes and the whisker in Figure 2. In the former case, it means that by sampling the weights the scores tend to give similar results in each run, which means that the criteria have similar performances. In the latter case, it means that for each sampled weighting profile more differentiated scores are calculated, which means that there are more divergent criteria performances within a country. It is important to note that this effect is relatively stronger for countries in which the effect of the divergent criteria performance is minimal (e.g., Switzerland, Albania), while it has no or relative minimal effects for countries (e.g., Portugal, Cyprus, Greece) in which the divergent criteria performances are large. This shows how different weight profiles could help in depicting country scores variation when the divergent criteria performance is minimal. Finally, worth noting is that generally the means are not the same. In fact, the mean of the ESRI case “sampled weights” looks slightly higher or lower than the ESRI base case scenario. The latter is most probably caused by the increased number of resulting scores in the first case, which improves the robustness of the results.

4.2. Deterministic vs. Uncertain Criteria

Figure 2b presents the comparison between the ESRI base case scenario and the ESRI "uncertain criteria". As in section 4.1, the countries are classified in descending order according to the ESRI mean score of the base case scenario. This comparison reveals differences in scoring results for most of the countries. The top-performer (Switzerland) and the low-performer (Albania), where the effect of having many top-performing criteria and low-performing criteria, respectively, show a relative low effect of the criteria uncertainty on their final rankings. This is also valid for some of the average-performing countries, e.g., Norway, Italy, and Estonia, where the comparison between the two versions of ESRI show a similar result, indicating that the effect of the criteria uncertainty is limited, due to their lower volatility. On the other hand, the effects of criteria uncertainty on the scoring and rankings of the other countries are significant, showing an average score decrease (e.g., Switzerland, UK, Ireland, Netherlands, Belgium, Poland) or an average score increase (e.g., Island, Hungary, Serbia, Greece, Bosnia & Herzegovina, Albania), indicating how strong is the effect on the scores of countries with criteria affected by large volatility, i.e., large standard deviation.



-

b)

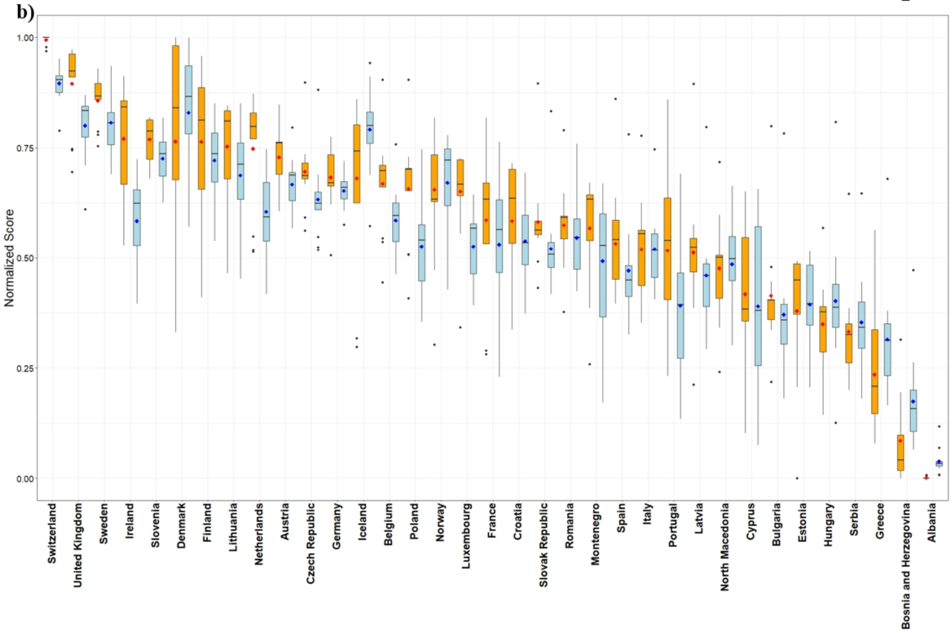


Fig. 2. Boxplot of normalized resilience scores per country. For each country, the red/blue square is the mean score value of the Electricity Supply Resilience Index (ESRI), while the black horizontal line represents the median. A box contains 50% of all scores, that is, it vertically extends from the first quartile to the third quartile of all scores. The whiskers extend to 3/2 of the length of the box. The black points represent outperformers. a) Results for the ESRI base case scenario (orange) and the ESRI "sampled weights" case (light blue). b) Results for the ESRI base case scenario (orange) and the ESRI "uncertain criteria" case (light blue).

5. Conclusions

This study presents an extension of the Electricity Supply Resilience Index (ESRI), developed to benchmark countries' electricity supply resilience. The ESRI is a composite indicator (CI) constructed according to the widely accepted ten-step methodological framework for good practice in the construction of CIs (Organisation for Economic Co-operation and Development (OECD), 2008). In the past, ESRI has focused on analyzing the sensitivity of the scoring/ranking of countries to the combination of normalization/aggregation functions, used to score the index (Gasser et al., 2020). In this context, two additional robustness analyses are considered in addition to the ESRI sensitivity analysis. First, the impact of different criteria weighting profiles on the final scoring/ranking of countries. Second, the impact of the presence of criteria uncertainty on the final scoring/ranking of countries. In this study, the ESRI extension was applied to its 17-criteria version for 35 European countries affiliated to ENTSO-E (Siskos and Burgherr, 2022). Direct ranking comparisons between previous studies are not possible because of differences in countries and/or numbers of criteria and criteria values and time periods (Gasser et al., 2020; Lindén et al., 2021; Siskos and Burgherr, 2022). In this study, we compare a base case scenario with other scenarios calculated by using the same information.

The comparison of the ESRI base case scenario with the ESRI "sampled weights" shows a generally limited impact of the weights on the final scoring/ranking of countries, which indicates that even under diverging preferences the scores/ranks are stable. However, the results indicate that different weighting profiles could help to represent the variation in the countries' scores when the divergent criteria performance is minimal, which is not the case when the latter is significantly large and therefore embedded in the combination of normalization/aggregation functions.

On the other hand, the ESRI "uncertain criteria" highlights that the differences are large or small for countries described by criteria with high or low volatility, i.e., standard deviation. Furthermore, in contrast to the previous comparison, where the score result is stable, the scoring for some of the countries varies by either increasing or reducing the scores, showing how strong a criterion uncertainty could be on the final scoring/ranking of a country.

Future research should focus on collecting additional data for some of the criteria considered, either to increase the robustness of the uncertainty criteria or to introduce uncertainty in the deterministic criteria that were kept this way due to the lack of information found by the authors. Furthermore, it would be of interest understanding what drives the scoring/ranking stability for a country, both in terms of changes in weights and changes in criteria values. In this context, a further development of the ProMCDA tool could be the ability to extract weightings and/or years in which a country performs best/worst.

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. Peter Burgherr acknowledges the support of the Swiss Federal Office of Energy as part of the SWEET consortium SURE. The authors bear sole responsibility for the conclusions and the results.

References

Bozdogan, H., 1987. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. Psychometrika 52, 345–370. https://doi.org/10.1007/BF02294361

Catalli, F., Spada, M., 2023. ProMCDA - Probabilistic Multi Criteria Decision Analysis.

Cimellaro, G.P., 2018. Urban Resilience for Emergency Response and Recovery. Springer, Cham, Switzerland.

Cinelli, M., Coles, S.R., Kirwan, K., 2014. Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment. Ecol. Indic. 46, 138–148. http://dx.doi.org/10.1016/j.ecolind.2014.06.011

Cinelli, M., Spada, M., Kim, W., Zhang, Y., Burgherr, P., 2020. MCDA Index Tool: an interactive software to develop indices and rankings. Environ. Syst. Decis. https://doi.org/10.1007/s10669-020-09784-x

Dobbie, M.J., Dail, D., 2013. Robustness and sensitivity of weighting and aggregation in constructing composite indices. Ecol. Indic. 29, 270–277. https://doi.org/10.1016/j.ecolind.2012.12.025

El Gibari, S., Gómez, T., Ruiz, F., 2018. Building composite indicators using multicriteria methods: a review. J. Bus. Econ. https://doi.org/10.1007/s11573-018-0902-z

ENTSO-E, 2023. European Resource Adequacy Assessment - Annex 3: Detailed Results. ENTSO-E.

ENTSO-E, 2017. Mid-Term Adequacy Forecast. Brussels, Belgium.

ENTSO-E, n.d. Power Statistic.

Eurostat, n.d. Eurostat Database.

Gasser, P., Lustenberger, P., Cinelli, M., Kim, W., Spada, M., Burgherr, P., Hirschberg, S., Stojadinovic, B., Sun, T.Y., 2019. A review on resilience assessment of energy systems. Sustain. Resilient Infrastruct. 1–27. http://dx.doi.org/10.1080/23789689.2019.1610600

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. Ecol. Indic. 110. https://dx.doi.org/10.1016/j.ecolind.2019.105731

Groothuis-Oudshoorn, C.G.M., Broekhuizen, H., van Til, J., 2017. Dealing with Uncertainty in the Analysis and Reporting of MCDA, in: Marsh, K., Goetghebeur, M., Thokala, P., Baltussen, R. (Eds.), Multi-Criteria Decision Analysis to Support Healthcare Decisions. Springer International Publishing, Cham, pp. 67–85. https://doi.org/10.1007/978-3-319-47540-0_5

Heinimann, H.R., Hatfield, K., 2017. Infrastructure Resilience Assessment, Management and Governance – State and Perspectives, in: Linkov, I., Palma-Oliveira, J.M. (Eds.), Resilience and Risk. Springer Netherlands, Dordrecht, pp. 147–187.

IEA, 2023. Energy Efficiency 2023. International Energy Agency, Paris, France.

IEA, n.d. IEA Electricity Information Statistics.

Kröger, W., 2017. Securing the Operation of Socially Critical Systems from an Engineering Perspective: New Challenges, Enhanced Tools and Novel Concepts. Eur. J. Secur. Res. 2, 39–55. https://doi.org/10.1007/s41125-017-0013-9

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. Environ. Model. Softw. https://doi.org/10.1016/j.envsoft.2021.105208

Mazziotta, M., Pareto, A., 2013. Methods for Constructing Composite Indices: One for All or All for One? Riv. Ital. Econ. Demogr. E Stat. 67.

Narula, K., Reddy, B.S., 2015. Three blind men and an elephant: The case of energy indices to measure energy security and energy sustainability. Energy 80, 148–158. http://dx.doi.org/10.1016/j.energy.2014.11.055

Organisation for Economic Co-operation and Development (OECD), 2008. Handbook on Constructing Composite Indicators: Methodology and User Guide. Organisation for Economic Co-operation and Development (OECD) Joint Research Center (JRC) European Commission.

Pohekar, S.D., Ramachandran, M., 2004. Application of multi-criteria decision making to sustainable energy planning—A review. Renew. Sustain. Energy Rev. 8, 365–381. https://doi.org/10.1016/j.rser.2003.12.007

Siskos, E., Askounis, D., Psarras, J., 2014. Multicriteria decision support for global e-government evaluation. Omega 46, 51–63. https://doi.org/10.1016/j.omega.2014.02.001

Siskos, E., Burgherr, P., 2022. Multicriteria decision support for the evaluation of electricity supply resilience: Exploration of interacting criteria. Eur. J. Oper. Res. 298, 611–626. https://doi.org/10.1016/j.ejor.2021.07.026

Streiner, D.L., 2003. Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. J. Pers. Assess. 80, 99–103. https://doi.org/10.1207/S15327752JPA8001_18

Swiss Re Institute, n.d. Sigma Explorer.

Valdés, J., 2018. Arbitrariness in Multidimensional Energy Security Indicators. Ecol. Econ. 145, 263–273. http://dx.doi.org/10.1016/j.ecolecon.2017.09.002.

Wu, G., Liu, L.-C., Han, Z.-Y., Wei, Y.-M., 2012. Climate protection and China's energy security: Win–win or tradeoff. Appl. Energy 97, 157–163. https://doi.org/10.1016/j.apenergy.2011.11.061.