Comprehensive resilience assessment of electricity supply security for 140 countries

Journal Article

Author(s):
Gasser, Patrick (ID); Suter, Johannes; Cinelli, Marco; Spada, Matteo; Burgherr, Peter; Hirschberg, Stefan; Kadziński, Miłosz; Stojadinovic, Bozidar (ID)
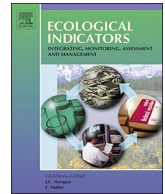
Comprehensive resilience assessment of electricity supply security for 140 countries

Patrick Gasser[a,*], Johannes Suter[a,b,c], Marco Cinelli[a,1], Matteo Spada[c], Peter Burgherr[c], Stefan Hirschberg[c], Miłosz Kadziński[d], Božidar Stojadinović[e]

[a] Future Resilient Systems (FRS), Singapore-ETH Centre (SEC), Swiss Federal Institute of Technology (ETH) Zürich, Singapore
[b] Group for Sustainability and Technology, Department of Management, Technology and Economics, Swiss Federal Institute of Technology (ETH) Zürich, Zurich, Switzerland
[c] Laboratory for Energy Systems Analysis (LEA), Paul Scherrer Institut (PSI), Villigen PSI, Switzerland
[d] Institute of Computing Science, Poznań University of Technology, Poznań, Poland
[e] Institute of Structural Engineering, Department of Civil, Environmental and Geomatic Engineering, Swiss Federal Institute of Technology (ETH) Zürich, Zurich, Switzerland

ABSTRACT

Indicator-based approaches are suitable to assess multi-dimensional problems. In order to compare a set of alternatives, one strategy is to normalize individual indicators to a common scale and aggregate them into a comprehensive score. This study proposes the Electricity Supply Resilience Index (ESRI), which is a measure of a nation's electricity supply resilience. Starting from an initial set of individual indicators derived through a structured selection process, the ESRI is calculated for 140 countries worldwide. To account for robustness of the resulting resilience index, 38 combinations of eight normalization methods and six aggregation functions were considered. Results show a clear country ranking trend, with robust top- and low-performing countries across all combinations. However, the ranking disparity becomes large for average performing countries, especially if their indicators show high variability. Furthermore, the differences of the rankings are quantified through the Rank Difference Measure (RDM), which identifies the categorical scales and the minimum aggregator as the most different ones. Finally, the effects of different compensation levels of the aggregation functions are discussed. The findings of the present study aim to provide recommendations for policymakers on how composite indexes results depend on assumptions and chosen approaches.

1. Introduction

Resilience is a multi-dimensional concept that is receiving growing attention in various disciplines, with many definitions and quantification methods proposed so far (Hosseini et al., 2016; Häring et al., 2017; Sharifi and Yamagata, 2016; Ouyang, 2014; Cimellaro, 2016; Francis and Bekera, 2014; Willis and Loa, 2015; Bergström et al., 2015). Even though there is no overall consensus, it is still widely accepted that a comprehensive resilience framework comprises both disruptive and recovery elements (Cimellaro, 2016). Within this paper, the resilience framework considered is the one developed by the Future Resilient Systems (FRS) program at the Singapore-ETH Centre (SEC) (Heinimann and Hatfield, 2017). This framework is generally applicable to infrastructure systems and comprises four dimensions:

(1) Resist: represents the system's ability to withstand disturbances within acceptable degradation levels.
(2) Restabilize: illustrates the ability to limit a performance decrease and re-establish key functionalities.
(3) Rebuild: describes the recovery process of system's performance back to normal.
(4) Reconfigure: characterizes the changes of the biophysical architecture/topology of the system to make it more fault-tolerant.

Within the energy sector, security of supply is an often-analyzed topic (Scheepers et al., 2007; Winzer, 2012; Laldjebaev et al., 2018; Kruyt et al., 2009; Sovacool and Mukherjee, 2011; Martchamadol and Kumar, 2013; Azzuni and Breyer, 2017), but research in the context of a resilience perspective is sparse. However, as the International Energy Agency (IEA) defines energy security as "the uninterrupted availability

of energy sources at an affordable price" (International Energy Agency, 2014), many of the security of supply indicators can also be used in a resilience context (Gasser et al., 2017). In fact, the "uninterrupted availability", in other words the length and severity of a disruption, is directly related to resilience as defined above. Furthermore, Gasser et al. (2017) demonstrate that the two concepts are closely related, as they both "aim at minimizing the frequency and severity of disruptions" (Gasser et al., 2017; Roege et al., 2014). Nevertheless, resilience extends the definition of energy security beyond frequency and severity, by accounting for the abilities to recover quickly and adapt to new operating conditions.

Due to the multi-dimensionality of energy security and resilience (Francis and Bekera, 2014; Kruyt et al., 2009; Sovacool and Mukherjee, 2011; Sovacool and Brown, 2010; Ang et al., 2015a), Multi-Criteria Decision Aiding (MCDA) is a promising approach as it provides a structured and justifiable process to develop a comprehensive assessment of the alternatives (e.g. countries, technologies, supply routes, type of equipment, etc.) under study (Pohekar and Ramachandran, 2004; Wang et al., 2009; Huang et al., 2011; Cinelli, 2017). Through MCDA, it is possible to aggregate several indicators representing the various dimensions into a comprehensive score, usually called an index or composite indicator (Joint Research Centre of the European Commission, 2008; Saisana and Saltelli, 2011; El Gibari et al., 2018). This allows for effective comparison between the alternatives, identification of benchmarks to be followed and areas requiring improvement (Sovacool, 2012). However, many studies using indices are also criticized for: (i) their simplicity and limited transparency as the final results are single numbers, (ii) their lack of accounting for uncertainties, and (iii) their lack of robustness as a measure to quantify the stability of the rankings or sensitivity analysis (Saltelli, 2007; Saisana et al., 2005; Grupp and Mogee, 2004; Burgass et al., 2017; Dobbie and Dail, 2013). In order to overcome these drawbacks, one option is to apply several MCDA methods so that the resulting multiple rankings can be compared and trends identified (Cinelli et al., 2014; Valdés, 2018; Greco et al., 2016; Velasquez and Hester, 2013; Ishizaka and Nemery, 2013; Dodgson et al., 2009).

Numerous indices exist to compare country performances based on economic, political, social or environmental measures (e.g. (Füssel, 2010; Brown and Matlock, 2011; Fowler and Hope, 2007; Böhringer and Jochem, 2007; Bandura, 2008). In fact, all of these studies are multi-dimensional and the country scores achieved are considered practical for policymakers because they are easy to understand and represent effective communication tools (Freudenberg, 2003). Nevertheless, to avoid misinterpretations, it is of crucial importance to follow a systematic index construction methodology and to be aware of the potential weaknesses involved (Joint Research Centre of the European Commission, 2008; Freudenberg, 2003). This is also true for the energy sector, where many indices measuring the energy security performance of countries have been published (Sovacool and Brown, 2010; Ang et al., 2015b; Angelis-Dimakis et al., 2012; Antanasijević et al., 2017; Augutis et al., 2009; Augutis et al., 2011; Augutis et al., 2012; Badea et al., 2011; Blyth and Lefevre, 2004; Boccauthor and Hanna, 2016; Bompard et al., 2017; Brown et al., 2014; Cabalu and Alfonso, 2013; Cabalu, 2010; Centre for Environmental Law Policy, 2018; Cohen et al., 2011; Doukas et al., 2012; Dunn et al., 2012; Eckle et al., 2011; Ediger and Berk, 2011; Erahman et al., 2016; Frondel, 2008; Geng and Ji, 2014; Glynn et al., 2017; Gnansounou, 2008; Gupta, 2008; Hu and Kao, 2007; Hughes and Shupe, 2010; Iddrisu and Bhattacharyya, 2015; Institute for 21st Century Energy, 2016; Institute for 21st Century Energy, 2017; Kamsamrong and Sorapipatana, 2014; Kanchana et al., 2016; Le Coq and Paltseva, 2009; Lefèvre, 2010; Li et al., 2016; María Marín-Quemada and Muñoz-Delgado, 2011; Martchamadol and Kumar, 2014; Molyneaux et al., 2012; Obadi et al., 2017; Onamics, 2005; Prambudia and Nakano, 2012; Radovanović et al., 2017; Ramanathan, 2005; Roupas et al., 2009; Selvakkumaran and Limmeechokchai, 2013; Sharifuddin, 2014; Sheinbaum-Pardo et al., 2012; Song et al., 2013; Sovacool et al., 2011; Wang and Zhou, 2017; World Energy Council, 2011; World Energy Council, 2018; World Economic Forum, 2018; Wu et al., 2007; Wu et al., 2012; Yao and Chang, 2014; Zeng et al., 2017; Zhang et al., 2011; Zhang et al., 2013; Zhang et al., 2017; Zhou and Ang, 2008). For example, as today's national economies still heavily depend on oil (British Petroleum, 2017), energy security indicators often describe vulnerabilities to oil supply disruptions, such as the oil vulnerability indices presented by Gupta (2008) and Roupas et al. (2009). For oil, gas and coal, Le Coq and Paltseva (2009) proposed a "risky external energy supply (REES) index", Blyth and Lefevre (2004) analyzed geopolitical proxy measures, Sato et al. (2017) applied the Shannon index to supplier portfolios, and Lefèvre (2010) constructed the Energy Security Price Index (ESPI) based on market concentration for fossil fuels and the Energy Security Physical Availability Index (ESPAI) based on supply flexibility. For the vulnerability of electricity sources in general, Gnansounou (2008) developed an electricity supply vulnerability index according to three sub-dimensions: (i) the net import of electricity, (ii) the concentration of import origins and risk that a dominant technology in an electricity generation portfolio is not accepted by the public, and (iii) the non-diversification of electricity generation. The Energy research Centre of the Netherlands created the Supply/Demand Index, which aims to assess energy security in the medium and long term (Scheepers et al., 2007). It covers primary energy sources, conversion and transport, and final energy demand. Glynn et al. (2017) applied a modification of this index to study the decarbonisation of the Irish energy system. In contrast, the IEA developed a Model of Short-term Energy Security (MOSES) (Jewell, 2011). Further indices available in the literature are, among others, the Energy Security Indices from the Asia Pacific Energy Research Centre (APERC) (Asia Pacific Energy Research Centre (APERC), 2007), the Energy Affinity Index about international relations (María Marín-Quemada and Muñoz-Delgado, 2011), the Energy Dependence Index (EDI) as the root mean square of the Energy Import Dependency (EID) and Energy Export Dependency (EED) (Kanchana et al., 2016), the Energy Indicators for Sustainable Development (EISD) (Vera and Langlois, 2007), the Aggregated Energy Security Performance Indicator (AESPI) (Martchamadol and Kumar, 2013), and the Sustainable Energy Development Index (SEDI) (Iddrisu and Bhattacharyya, 2015).

After having developed a relevant indicator set that usually includes indicators related to the diversity of generation, diversity of supply, energy intensity, fuel reserves, shares of renewables, efficiencies, greenhouse gases emissions, self-sufficiency and energy prices, the typical methodology consists in selecting a normalization method, weighting profile and aggregation function in order to calculate the final scores representing countries' energy security performance. There exists a plethora of ways to do so. For details on each one of them, the reader is referred to the handbook on constructing indices of the Joint Research Centre of the European Commission (2008) or Greco et al. (2016).

Three main research gaps can be identified from the available literature. First, most studies only incorporate elements related to the ability to resist disruptions and/or vulnerabilities caused by disruptive events. Therefore, there is a lack of information about indicators to quantify the abilities to rebuild and potentially reconfigure a system to make it more resilient towards future hazardous events. Second, a verification of the indicator selection process, reliability and suitability of the data set is usually missing. Third, only few studies build indices using more than one combination of normalization, weighting and aggregation method (Augutis et al., 2012; Boccauthor and Hanna, 2016; Gnansounou, 2008; Li et al., 2016; Ramanathan, 2005; Zeng et al., 2017; Zhang et al., 2017; Zhou and Ang, 2008). Knowing that each method has its pros and cons (Joint Research Centre of the European Commission, 2008; Gan et al., 2017), and especially that the rankings might vary to some extent (Narula and Reddy, 2015), there is a considerable lack of systematically assessing the robustness of the ranking.

Based on these premises, the objectives of this paper are threefold:

1. Development of an indicator set that comprehensively covers the various dimensions of resilience (i.e., resist, restabilize, rebuild and reconfigure) and allows evaluating the electricity supply resilience of countries.
2. Verification of the selection of the indicators by means of statistical coherence tests and assessment of their suitability to build indices.
3. Development of indices to assess the electricity supply resilience of 140 countries worldwide by accounting for multiple preferences of stakeholders. These preferences are represented by the consideration of 38 different approaches, in order to assess the robustness of the results. For this purpose, the Electricity Supply Resilience Index (ESRI) is proposed as a comprehensive and integrated measure, and it can be used to rank the countries.

This research is of interest for decision-makers as the methodology presented here can help them achieving an insightful understanding of the problem under consideration. Furthermore, indices are a promising approach for policymaking as they allow to position the country of interest with respect to the others, to identify successful strategies, to compare oneself with its benchmarks and to identify target countries to learn from. This is of interest for any governmental agency, research institute, university and company. As the index' construction methodology can affect the rankings due to different levels of compensation between indicators, the interested users can understand the consequences and resulting implications of the methodological choices. Compared to a single score based on a single normalization method and aggregation function, a robustness analysis enhances the credibility in the results, because it allows to study the stability of the rankings.

Section 2 starts with a general overview of the structured methodology applied to construct an index measuring countries' electricity supply resilience. Then the country and indicator selection process are detailed, the multivariate analyses are described, and the normalization and aggregation methods considered are introduced. Section 3 presents the country ranking results, points out general trends, and discusses the effects of compensation through different normalization and aggregation methods on the final rankings. Furthermore, the differences in the rankings are discussed through the RDM. It has to be noted that, even though an electricity supply resilience country ranking is provided hereby, this paper does not directly address strategies to improve the index. In fact, the value of this paper lies in the development of the novel indicator set, robustness analysis and discussion of how index construction methodologies might affect rankings. In Section 4, the main conclusions are given and directions for future research are proposed.

2. Methodology

A comprehensive MCDA application requires several steps (Belton et al., 2002). The methodology used in this paper is built upon the one proposed in the Handbook on Constructing Composite Indicators published by the Joint Research Centre (JRC) of the European Commission (EC) (Joint Research Centre of the European Commission, 2008). A detailed flow chart of the various steps of the methodology applied to the present case study is illustrated in Fig. 1, and presented in detail in the remainder of this section. In summary, the 15 steps of Fig. 1 represent:

1. Steps 1.1 and 1.2: Literature reviews in order to obtain a broad understanding of resilience and index construction methodologies (see Section 2.1).
2. Step 2: Development of a theoretical framework aimed at defining the research topic (see Section 2.1). The theoretical framework facilitates the subsequent steps of identifying indicators.
3. Step 3: Selection of an initial set of indicators (see Table 1) that is a

direct outcome of the literature review on electricity supply resilience indicators by considering the theoretical framework developed in step 2.
4. Steps 4.1 and 4.2: These steps represent the simultaneous selection of countries and assessment of the indicators based on four criteria: relevance, credibility of the data, accessibility of the data, and applicability and comparability between countries (see Section 2.2).
5. Step 5: Result of the indicator assessment conducted in step 4 (see Section 2.2).
6. Step 6: Data treatment in which outperformers are trimmed and missing values imputed (see Section 2.3).
7. Step 7: Multivariate analysis (indicator set verification via tests of statistical coherence) is performed in order to assess the structure of the data set (see Section 2.4).
8. Step 8: Verification of the results of the multivariate analysis (see Section 2.4) to determine if indices can be constructed. Otherwise, the country or indicator selection needs to be revised.
9. Step 9: The final indicator set that fulfills the multivariate analysis is obtained (see Table 2).
10. Step 10: Selection of data normalization methods (see Table 4). The normalization methods bring the indicators to a common scale, as the indicators are expressed in different units. A total of eight methods were considered: four ordinal ones (rank, percentile rank and categorical), three linear ones (standardized, min–max and target) and the logistic function.
11. Step 11: Selection of indicator weighting schemes (see Section 2.5). The weights represent the relative importance of the indicators towards the index. In order to study the effect of the normalization methods and aggregation functions solely, an equal weighting profile was considered.
12. Step 12: Selection of data aggregation functions (see Table 5). The aggregation functions combine the weighted indicators into an index. A total of six aggregation functions were considered: additive, geometric, harmonic, minimum, median and Condorcet.
13. Step 13: Selection of combinations of normalization methods, indicator weighting schemes and aggregation functions to construct indices (see Table 6). The 38 combinations considered represent an assessment of robustness by accounting for several index construction methodologies.
14. Step 14: Comparison of scores and rankings (see Section 3). The results are discussed in relation to three ranking comparison measures: (1) variability assessment through a boxplot, (2) Rank Difference Measure (RDM) and (3) Rank Acceptability Indices (RAIs). In particular, the effects of normalization methods and aggregation functions on the indices are analyzed.
15. Step 15: Interpretation of results and recommendations for policymakers (see Section 3).

2.1. Resilience features

Initially, an extensive literature review on electricity supply resilience (step 1.1 in Fig. 1 (Gasser et al., 2017)) was carried out in order to identify how each case study was performed. In parallel, the recognized procedures for constructing indices were studied (step 1.2 in Fig. 1). Given the fact that a plethora of indicators are reported in the literature, developing a theoretical framework that gives a clear understanding and structure of the multidimensional phenomenon to be measured, supports the indicator selection process (Joint Research Centre of the European Commission, 2008). Therefore, relevant features are defined for each resilience dimension (step 2 in Fig. 1, see Fig. 2). These descriptive features were used to establish the first set of measurable resilience indicators that cover all relevant topics in the four resilience dimensions (step 3 in Fig. 1). Detailed explanations of the resilience dimensions and the associated features are given in Section S1 of the Electronic Supplementary Information (ESI).
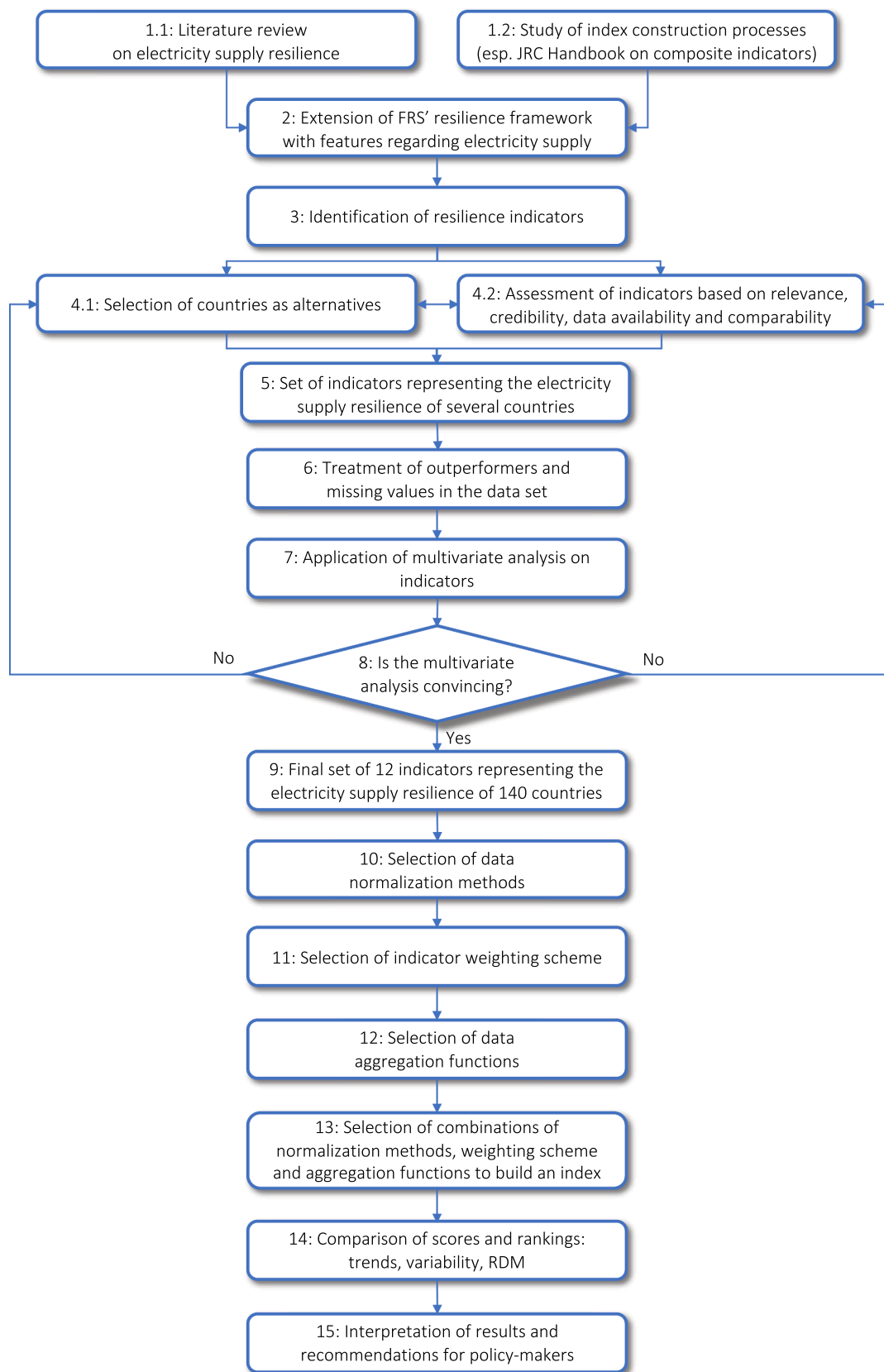
Fig. 1. Flow chart of the methodology followed in this study (adapted from Joint Research Centre of the European Commission (2008)).

Table 1

Overview of the four resilience dimensions, associated features and resilience indicators with units. The indicators were rated as fulfilling the four assessment criteria extensively (++), sufficiently (+) or insufficiently (−). The final set of the 12 indicators considered (bold ID numbers) are the ones scoring at least sufficiently (+) over all criteria. The ones not included in the final set did score insufficiently (−) in at least one of the criteria (italics, not numbered). Some indicators are represented twice, because they are related to multiple resilience dimensions. However, they are not double-counted in the index construction.

Resilience dimension	Feature	ID	Indicator name	Indicator unit	Criteria for indicator assessment				Data sources
					1	2	3	4	
Resist	Quality and reliability of the electricity system	1	System Average Interruption Duration Index (SAIDI)	Hours per year and customer	++	+	+	+	World Bank (2016)
		2	Severe accident risks	Fatalities/GWeyr[1]	++	++	+	+	International Energy (2015), Burgherr and Hirschberg (2014), International Renewable Energy Agency (2017)
	Stability of the general environment		*Investment into energy infrastructure*	*USD*	++	+	−	−	Asian Development Bank (2017)
		3	Control of corruption	% rank	+	++	++	+	World Bank (2015)
		4	Political stability and absence of violence/terrorism	% rank	+	++	++	+	World Bank (2015)
			Price volatility	*% change*	+	−	−	−	Sovacool et al. (2011)
Restabilize	Diversity of electricity generation	5	Electricity mix diversity	Normalized Shannon index	++	++	++	++	International Energy Agency (2015)
	Fuel security	6	Electricity import dependence	Ratio	++	++	++	++	International Energy Agency (2015)
			Spatial distribution of generation	*–*	++	−	−	−	Wender et al. (2017)
			Fuel reserves to consumption	*Years*	+	+	+	−	BP (2017)
			Diversity of fuel supply chains	*Normalized Shannon index*	++	−	−	−	Pettit et al. (2010)
	Control of system and cascades	7	Equivalent availability factor	%	++	+	+	+	International Energy Agency (2015), Volkart et al. (2016), U.S. Energy Information Administration (2018)
			Emergency preparedness measures	*–*	+	−	−	−	Cutter (2015)
Rebuild	Financial resources	8	GDP per capita	2010 USD per capita	++	++	++	++	International Energy Agency (2015)
		9	Insurance penetration	premiums paid in % of GDP	++	+	+	+	Swiss Re (2018)
			Private investment into energy infrastructure	*USD*	+	+	−	−	World Bank (2016)
	System governance and processes	4	Political stability and absence of violence/terrorism	% rank	+	++	++	+	World Bank (2015)
		10	Government effectiveness	% rank	++	++	++	+	World Bank (2015)
	Skilled personnel resources		*Engineers in the economy*	*%*	++	+	−	−	Hosseini and Barker (2016)
			Technical qualified personnel in energy sector and electric utilities	*%*	++	+	−	−	Hosseini and Barker (2016)
	Technical resources	11	Average outage time	Hours	++	+	+	+	World Bank (2016)
		7	Equivalent availability factor	%	++	++	++	++	(International Energy Agency (2015), International Renewable Energy Agency (2017) Volkart et al. (2016)
Reconfigure	System adaptability	5	Electricity mix diversity	Normalized Shannon index	++	++	++	++	International Energy Agency (2015)
			International agreements	*Normalized Shannon index*	++	−	−	−	–
	Integration of new technologies	12	Ease of doing business	Distance to frontier	+	+	+	+	World Bank (2016)

[1] GWeyr stands for GigaWatt$_{electric}$year. It is a unit of electrical energy. 1 GWeyr = $1/(8760 * 10^{6})$ k Wh of electricity.

Table 2

Performance matrix with selected countries and 12 indicators. The performance matrix including all 140 countries is available in Table S1 in the ESI. In grey are the inserted missing values. In orange are the outperformers trimmed to the lower bound and in blue are the ones trimmed to the upper bound. An upward pointing arrow (green) indicates better performance for higher values, whereas a downward pointing arrow (red) indicates better performance for lower values.

0 - Country	1 - SAIDI	2 - Severe accident risks	3 - Control of corruption	4 - Political stability and absence of violence/terrorism	5 - Electricity mix diversity	6 - Electricity import dependence	7 - Equivalent availability factor	8 - GDP per capita	9 - Insurance penetration	10 - Government effectiveness	11 - Average outage time	12 - Ease of doing business
Measurement unit	Hours per customer per year	Fatalities/GWeyr	Percentile rank	Percentile rank	Normalized Shannon index	Ratio	%	2010 USD per capita	%	Percentile rank	Hours	Distance to frontier (100 = best, 0 = worst)
Preference order of the values	↓	↓	↑	↑	↑	↓	↑	↑	↑	↑	↓	↑
Albania	21.4	7.03	38.46	58.10	0.00	1.02	39%	4,543	0.90%	53.85	2.15	66.
Cambodia	21.4	3.53	12.02	49.52	0.39	1.16	64%	1,021	0.30%	25.00	1.40	52.
China	1.3	7.03	48.56	26.19	0.42	0.95	74%	6,498	1.63%	68.27	3.25	63.
Congo, Dem. Rep.	6.8	7.01	8.17	4.76	0.01	0.81	39%	384	0.30%	2.88	1.71	34.
Denmark	0.4	0.04	98.08	76.67	0.57	1.14	54%	50,108	2.69%	97.60	0.80	83.
Eritrea	6.8	0.95	7.69	19.05	0.01	0.86	85%	528	0.40%	4.81	1.71	32.
Finland	0.2	0.03	99.52	87.14	0.71	1.16	73%	45,208	2.18%	96.15	0.50	80.
Iceland	0.5	0.00	95.67	95.71	0.25	0.97	53%	45,939	2.20%	90.87	1.00	78.
Iraq	21.4	0.97	4.81	2.86	0.31	0.64	83%	5,120	1.58%	9.62	2.33	44.
Italy	0.7	0.05	57.69	58.57	0.76	1.09	68%	33,912	2.06%	69.23	0.28	71.
Japan	0.4	0.08	91.35	89.05	0.65	0.96	78%	47,142	2.55%	95.19	4.00	75.
Libya	21.4	0.50	0.96	3.33	0.30	0.64	85%	5,447	0.40%	1.92	3.11	32.
Luxembourg	0.2	0.02	97.12	98.10	0.49	1.16	54%	50,108	1.79%	93.27	1.00	68.
Myanmar	6.8	4.20	20.67	10.48	0.33	0.84	58%	1,643	0.10%	10.10	1.71	38.
Nepal	6.8	7.02	32.21	14.29	0.01	1.12	39%	690	1.58%	12.98	1.71	59.
Netherlands	0.3	0.08	94.71	80.48	0.57	1.03	80%	50,108	3.89%	97.12	1.00	75.
New Zealand	2.4	0.02	100.00	99.05	0.56	0.94	56%	36,236	3.89%	98.56	1.71	86.
Niger	21.4	0.79	30.77	13.33	0.31	1.16	84%	384	0.70%	31.25	1.50	45.
Norway	1.8	0.00	99.04	91.43	0.10	0.84	40%	50,108	2.21%	98.08	0.82	82.
Paraguay	15.9	7.03	15.87	48.57	0.00	0.64	39%	3,822	1.20%	17.31	1.71	59.
Qatar	0.7	0.12	78.37	84.29	0.00	0.94	85%	50,108	1.50%	77.40	1.75	65.
Singapore	0.0	0.12	96.63	96.19	0.11	0.98	85%	50,108	1.69%	100.00	0.00	84.
South Korea	0.0	0.07	66.83	53.81	0.56	0.97	85%	25,021	3.89%	79.81	0.00	83.
South Sudan	6.8	0.95	0.48	2.38	0.02	0.94	85%	332	1.58%	0.48	1.71	35.
Spain	0.3	0.05	69.71	55.71	0.84	0.91	66%	30,486	2.75%	85.10	0.50	73.
Sweden	1.9	0.01	98.56	80.95	0.53	0.82	58%	50,108	1.88%	96.63	1.46	80.
Switzerland	0.1	0.01	97.60	96.67	0.41	0.92	58%	50,108	3.89%	99.52	1.00	75.
Syrian Arab Republic	6.8	0.52	1.92	0.00	0.31	0.84	84%	919	0.30%	5.29	1.71	41.
Togo	6.8	5.10	25.48	38.10	0.34	1.16	53%	554	1.10%	11.06	1.71	46.
Turkmenistan	6.8	0.12	5.77	42.86	0.00	0.73	85%	6,937	1.58%	19.23	1.71	62.
United Kingdom	0.4	0.05	93.75	61.43	0.75	0.98	76%	41,196	2.44%	94.23	2.00	82.
United States of America	0.6	0.07	89.90	67.14	0.67	0.96	79%	50,108	3.89%	89.90	2.00	82.
Vietnam	21.4	2.79	41.83	50.00	0.49	0.92	68%	1,685	0.74%	55.29	1.98	59.
Yemen	6.8	0.62	3.37	0.48	0.29	0.74	85%	775	0.20%	3.37	1.71	44.

2.2. Selection of countries and indicators

The operationalization of the resilience features was enabled by the development of indicators in conjunction with the countries to be analyzed in the subsequent case study. Consequently, the choice of the countries and indicators are interrelated (steps 4.1 and 4.2 in Fig. 1). The aim is to develop a consistent, credible and quantifiable indicator set that covers all resilience dimensions and can be applied to a broad selection of countries to identify trends and draw general conclusions. The final sets of countries and indicators are presented in Sections 2.2.1 and 2.2.2, respectively.

2.2.1. Country set selection

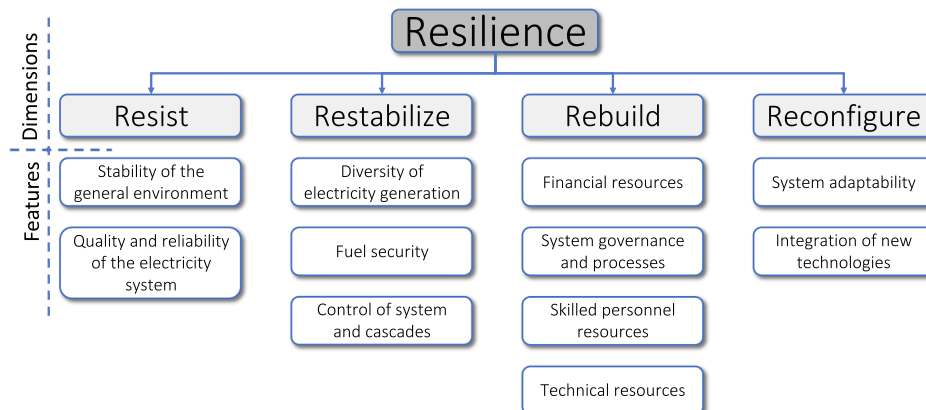The list of the World Bank's Worldwide Governance Indicators

Fig. 2. Extension of the FRS' resilience framework with features relevant in an electricity supply perspective (adapted and extended from Gasser et al. (2017)).

(WGI) consisting of 214 countries or territories was used as a starting point (World Bank, 2015). Based on the availability of energy-related data retrieved from the International Energy Agency (IEA) (International Energy Agency, 2015), the final set includes 140 countries (step 4.1 in Fig. 1). These countries cover all continents, represent more than 96% of the world's population and 99.6% of the world's electricity consumption.

2.2.2. Indicator set selection

The indicators selection process started from the work published by Gasser et al. (2017) (step 1.1 in Fig. 1). Web of Science and Google Scholar were used to identify the most appropriate literature. Multiple types of documents, such as peer-reviewed research articles, conference papers, technical and policy reports, books and thesis reports were considered. The search was conducted on the basis of keyword combinations, including resilience, security of supply, energy security, electricity, indicator, and multi-criteria decision analysis. The category "research articles" also includes recent review papers. Hence, the resilience-related indicators identified through this process are deemed to be comprehensive and suitable to measure the different dimensions of resilience.

After defining the scope of the case study, a smaller set was retained (step 3 in Fig. 1) and studied in detail. This set is presented in Table 1, and each indicator was assessed according to four criteria (step 4.2 in Fig. 1) (Foxon et al., 2002; Jasiński et al., 2018):

1. Relevance: Is the indicator directly representing or linked to the corresponding resilience dimension and feature? As resilience is a broad term, many indicators could be loosely connected to it. However, based on the extensive literature review conducted and the theoretical framework developed, it was judged if some indicators provide more precise representations of resilience in general, or its specific dimensions and features (see Sections S1 and S2 of the ESI).
2. Credibility of the data: Are there credible data sources? This was assessed based on expert judgement, with widely recognized international institutions (e.g. the World Bank and the IEA) being assumed as more credible.
3. Accessibility of the data: Is data available for all countries?
4. Applicability and comparability: Is it possible to compare the data between countries? The more a single calculation methodology or data source was used to quantify an indicator for each country and the fewer discrepancies between countries, the more applicable and comparable the indicator becomes.

As a result, ten indicators of the reduced set were not considered anymore in the indicator set, mainly because of data availability and comparability issues. Details for these indicators can be found in Section S2 of the ESI. The indicator set with which further calculations were performed (step 5 in Fig. 1) consists of 12 indicators that cover all considered resilience dimensions and score at least sufficiently (+) on each of the four assessment criteria (Fig. 2 and Table 1)[2].

1. Indicator 1 (i_1): System average Interruption Duration index (SAIDI) (Resilience dimension (RD): resist)
   The System Average Interruption Duration Index (SAIDI) is a measure of the total duration of electricity supply interruptions per customer per year. It is a commonly used indicator and a direct representation of the quality and reliability of electricity supply

---

[2] It is important to note that some indicators can be assigned to more than one resilience dimension. This is not an issue when constructing the index if such an indicator is counted only once. However, if the aim was to score each of the resilience dimensions separately, the weight of this indicator would have to be spread over the corresponding dimensions.

(Layton, 2004). A high SAIDI value indicates that electricity supply disruptions are frequent, meaning that the system does not perform satisfactorily at resisting disturbances.

2. Indicator 2 (i_2): Severe accident risks (RD: resist)
   Severe accident risk quantifies the number of fatalities per unit of electricity produced. Initially, the indicator is quantified for each generation technology and then aggregated according to the production mix of each country. The fatalities do not only include events from the actual power production, but all stages of the production chain (Burgherr and Hirschberg, 2008; Hirschberg et al., 2004). The specific normalized fatality rates per technology and geographical region are given in Table S4. A high value indicates that severe accidents are frequent resulting in a higher fatality rate. This generally happens in countries that have lower safety standards or lower technological know-how (Burgherr et al., 2011; Burgherr et al., 2012).

3. Indicator 3 (i_3): Control of corruption (RD: resist)
   Control of corruption is a governance measure that captures to which extent public power is used for private gains (World Bank, 2015). Instead of trying to improve the existing infrastructure for the community as a whole, private interests have higher values in more corrupt environments. Hence, higher corruption is more likely to lead to disruptions within the system, because the individual components are less robust and there is usually a considerable lack of well-established processes during critical situations (Wang and Zhou, 2017). Such processes are of crucial importance so that employees in the electricity sector know how to react rapidly during emergency crises. Furthermore, information about the real status of the system may be inscrutable or misleading (Sovacool and Mukherjee, 2011), which leads to high inefficiencies (Gulati, 2006). Finally, electricity theft is more common in corrupt environments, making the regular operation of the electricity grid more unstable, which ultimately results in more frequent disruptions (Transmission Distribution World, 2025).

4. Indicator 4 (i_4): Political stability and absence of violence/terrorism (RD: Resist and rebuild)
   Similar to the control of corruption, a higher political stability and absence of violence or terrorism leads to fewer disruptions in the electricity system (Ang et al., 2015a). In fact, as political stability fosters long-term investments, it can safely be assumed that the infrastructure in politically stable regions is of higher quality (Wang and Zhou, 2017). Major projects within the electricity sector often require large efforts, and in some cases, such projects are realized only after several years or decades. Furthermore, since electricity supply is a critical infrastructure for a community, it is a target for terrorism and violent armed groups leading to higher likelihoods of potential supply disruptions. Finally, energy security and political stability are interlinked, as "access to stable sources of energy is one prerequisite for state stability" (Organization for Security).

5. Indicator 5 (i_5): Electricity mix diversity (RD: Restabilize and reconfigure)
   Diversity is one of the key features for the ability to restabilize a system. In the case of unforeseen supply disruptions, it allows an easier shift from one technology to another or to modify the supply routes (Kruyt et al., 2009). Hence, a diverse supply is "a good way to hedge against unforeseen supply risks" (Molyneaux et al., 2012). Diversity is also a way to mitigate the effects of technology lock-in (Sovacool, 2010). It is calculated using the normalized Shannon index (Spellberg and Fedor, 2003):

$$H' = -\frac{1}{\ln N} \sum_i p_i \cdot \ln p_i \tag{1}$$

where p_i is the share of technology i in the generation mix and N denotes the total number of technologies. In the present study, coal, oil, natural gas, biomass, nuclear, hydropower, geothermal, solar

photovoltaic, wind and other sources (e.g. solar thermal, tides, waves, ocean) were considered. In the present case, the Shannon index was normalized between 0 and 1 by dividing by natural logarithm of *N*. This allows for easier interpretation, as the index tends to 1 when the technologies have roughly equal proportions (Ramezani, 2012). On the opposite, a lower value indicates less diversity, with the extreme case of 0 for a country that relies on a single fuel source.

6. Indicator 6 (i_6): Electricity import dependence (RD: restabilize)
   The import dependence is defined as the ratio between the consumption and production. Therefore, a country with a value lower than 1 (production > consumption) has export capabilities. This represents more flexibility on the flows to better absorb disruptions, because of the higher ability to reroute flows (Sovacool and Mukherjee, 2011). If needed, the production excess could be used for its own consumption. On the other hand, a value higher than 1 (consumption > production) represents an import dependence, making a country vulnerable to shortages. Furthermore, a country reliant on imports is affected by the stability of the country supplying those imports, which potentially increases the chances of non-delivery (Molyneaux et al., 2012).

7. Indicator 7 (i_7): Equivalent availability factor (RD: Restabilize and rebuild)
   This indicator measures the availability of a plant to be controlled or dispatched due to partial or total outages caused by technical failures or resource limitations (Volkart et al., 2016). Partial or full plant outages may be due to either scheduled maintenance or forced outages. Partial outages are reflected by using an equivalent availability factor given by the available annual generation, divided by rated capacity times 8760 hours per year. Not all available plants are always used (dispatched), so the equivalent availability factor forms an upper bound to the plant capacity factor that can be achieved. Furthermore, it does not apply to a non-dispatchable technology (Hirschberg et al., 2008). Hence, for hydropower, wind and solar, the capacity factor was used instead, because these technologies generate as much electricity as possible, constrained by resource availability. The technology specific equivalent availability factors are given in Table S5. For the equivalent availability factor of a country, a weighted sum according to the country's generation mix is calculated. The probability that a country with a higher equivalent availability factor will be able to produce electricity when necessary is higher. This potentially results in smaller-scaled blackouts and faster recovery. Hence, the capacities to re-stabilize and rebuild supply after a disruptive event are enhanced.

8. Indicator 8 (i_8): GDP per capita (RD: rebuild)
   To assess the financial strength of a country and its economy, the GDP is a well-accepted and commonly used indicator. A financially strong country is in a better position to acquire the needed technical, material and human resources to rebuild a system after failure (Kruyt et al., 2009). Such a country can therefore expect a faster recovery.

9. Indicator 9 (i_9): Insurance penetration (RD: rebuild)
   It has been shown that insurance is the fastest and most equitable mean of financing reconstruction (Asgary et al., 2015). Furthermore, it decreases the workload for governments and shifts the administrative expenses to insurance companies in the private sector. Being properly insured provides faster access to the financial resources necessary to rebuild a system, making it a crucial element for the rebuild dimension of resilience. Furthermore, insurance "also serves as a market-based incentive mechanism to encourage investments in mitigation measures in return for reductions in insurance premiums" (Tonn et al., 2018).

10. Indicator 10 (i_{10}): Government effectiveness (RD: rebuild)
    Rebuilding damaged infrastructure is a complex and time intensive process involving many stakeholders. Government effectiveness represents the quality of public services and the quality of policy formulation and implementation. Therefore, an effective government with clearly defined processes and governance leads to faster recoveries, because stakeholders and employees know precisely what to do and how to do it (Heinimann and Hatfield, 2017). In summary, a low government effectiveness is a threat to energy security (Martišauskas et al., 2018).

11. Indicator 11 (i_{11}): Average outage time (RD: rebuild)
    The average outage time is the ratio between SAIDI and the System Average Interruption Frequency Index (SAIFI). Hence, it represents the average length of disruptions until successful recovery, which is directly linked to resilience' rebuild dimension. Countries with for example detailed emergency preparedness measures, defined processes, human and financial capabilities, are expected to recover faster (Finster et al., 2016).

12. Indicator 12 (i_{12}): Ease of doing business (RD: reconfigure)
    The ease of doing business index represents the conduciveness of the regulatory environment to start and operate businesses (World Bank, 2017). A country ranked high has faster and simpler regulations to implement technological change. Hence, this indicator can be used to measure the difficulty to reconfigure a system with more advanced components, technologies and monitoring equipment (European Commission, 2017). Furthermore, private sector involvement would face difficulties in a country with an unfavorable environment to do business (Laldjebaev et al., 2018). Therefore, substantial efforts would be required to adapt to new conditions, leading to less capabilities of reconfiguration. Finally, the ease of doing business index is linked to investment risk (Yan et al., 2017). Countries in need for the long-term investments that energy infrastructures require would thus encounter more difficulties in getting them if their environment is not business-friendly.

2.3. Data treatment

The data for the 12 indicators for each of the 140 countries was obtained from the World Bank, IEA, Swiss Reinsurance Company Ltd. (Swiss Re), Paul Scherrer Institute (PSI), the U.S. Energy Information Administration and the International Renewable Energy Agency (IRENA). The details are given in Table 1. Once the data set was constituted, the next step in the construction of indices was the treatment of outperformers[3] and missing data (step 6 in Fig. 1) (Saisana and Saltelli, 2011).

Outperformers are extreme observations with respect to the other values of an indicator. They can have a strong impact on the final result, depending on which normalization method is applied (Joint Research Centre of the European Commission, 2008; Hawkins, 1980). To avoid such issues, outperformers were trimmed. They were identified with the Interquartile Range (IQR) method, which does not assume normality of the data (Seo, 2006). Values are considered as outperformers if they lay outside 1.5 times the IQR from the first and third quartiles (Q1 and Q3 respectively) (Ghasemi and Zahediasl, 2012). In mathematical terms, x_i is an outperformer if $x_i < Q1 - 1.5 * IQR$ or $x_i > Q3 + 1.5 * IQR$. As shown in Table S3, this method identifies 69 outperformers out of 1680 values (4%) in the present data set. These were trimmed to the nearest value that is not an outperformer. To study the consequences of trimming the outperformers, the entire calculation and analysis that follows was also conducted for the untrimmed data set. Corresponding results are presented in Section S12.

Regarding missing data, indicators i_1, i_9, i_{11} and i_{12} have missing values. These were assigned to the means of the indicators' values in order to minimize data distortion (Joint Research Centre of the European Commission, 2008). It is important to note that the insertion

---

[3] The term "outperformer" is used instead of "outlier" to indicate that an extreme value is based on measurements and official statistics, thus simply representing a country's performance.

of the missing values was done after trimming the outperformers. Table 2 provides an overview of the performance matrix, which is used for the further steps of the present study. Due to the large size of the table, only selected countries a shown here, while the full table is available in Table S1. For indicators i_1, i_2, i_6 and i_{11}, a lower value indicates better performance, whereas for the rest of the indicators, a higher value is better. This is essential for the data normalization step, as all the indicators of the normalized data sets need to point in the same direction (e.g. a higher value indicates better performance). Therefore, the preference order for indicators i_1, i_2, i_6 and i_{11} was inverted during the normalization process.

2.4. Coherence of indicator set structure (multivariate analysis)

Multivariate analysis is performed to evaluate the reliability of an indicator set and its internal consistency to develop an index (steps 7 and 8 in Fig. 1) (Joint Research Centre of the European Commission, 2008). These qualities of the data set were measured using Cronbach's Alpha and correlations. All calculations were carried out with the Statistical Package for the Social Sciences (SPSS) (Field, 2013).

Cronbach's Alpha (Cronbach, 1951) is the most widely used index to assess the reliability of a scale (Streiner, 2003). It measures how closely related a set of indicators are as a group. Cronbach's Alpha values lower than 0.7 indicate questionable internal consistency and thus imply a need for further multivariate analysis on the indicators (Nunnally and Bernstein, 1978), whereas values higher than 0.9 indicate excessive redundancy among indicators (Streiner, 2003). The present case study data set has a Cronbach's Alpha of 0.84, which is within the desirable range for consistent composite scales developed for research purposes (0.8 to 0.9) (Streiner, 2003).

Correlation between the indicators was assessed through Spearman's rho, a nonparametric measure that does not assume data normality (Spearman, 1904). The main advantage of using Spearman's rho, being a rank correlation measure, is that it is insensitive to outliers (Mukaka, 2012). Similar to Cronbach's Alpha, a certain overall positive degree of correlation (0.3 to 0.9) is desirable as it shows the degree to which the indicators point at the same direction. However, extremely high correlations (i.e. greater than 0.9) are signs of redundancy and the corresponding indicators could potentially be removed from the data

set. Values between -0.3 and 0.3 indicate that there is no significant correlation between variables. Finally, on the negative side, a value between -0.3 and -0.5 indicates a weak negative correlation, and values lower than -0.5 a gradually stronger negative correlation. In the present case study, there are mostly positive correlations, indicating that constructing an index with the selected indicator set is suitable (see Table 3) (Becker et al., 2017). High correlations exist between indicators i_3, i_4, i_8, i_{10} and i_{12}, with the highest one (0.94) found between i_3 and i_{10}. Nevertheless, as the mean correlation coefficients of indicators i_3 and i_{10} with the other indicators in the data set (0.477 and 0.521, respectively) are acceptable, and because these two indicators represent different resilience dimensions (resist and rebuild, respectively), both were maintained in the dataset. Furthermore, indicators i_5, i_6, i_7 and i_{11} are rather uncorrelated to the other indicators. The electricity import dependence indicator i_6 has a slight negative correlation (-0.3 to -0.5) with most of the other indicators. Nevertheless, the indicator of electricity import dependence was not excluded because (i) the negative correlations are still weak (Mukaka, 2012), (ii) this indicator is central to the restabilize resilience dimension, and (iii) due to its very good Cronbach's Alpha value. There are reported cases where a strong negative correlation leads to the conclusion that indicators should not be aggregated, but these levels of negative correlations are much higher than the present ones (Saisana and Philippas, 2012). Looking at the average correlation coefficients, Table 3 also gives an overview of the three most positively correlated indicators and the three most negatively correlated ones. In summary, Cronbach's Alpha and Spearman's rho correlation analyses confirm that the indicator set is suitable to construct an aggregated index.

2.5. Normalization methods and aggregation functions

Based on the multivariate analysis in Section 2.4, the final set of 12 indicators was validated (step 9 in Fig. 1). The next three steps to construct indices are to normalize the data (step 10 in Fig. 1), select a weighting scheme (step 11 in Fig. 1) and aggregate the normalized dataset (step 12 in Fig. 1). Normalization brings all indicators to a common scale so that they can be compared with each other. The weighting scheme defines how much importance is assigned to each individual indicator. In the present study, the aim is to study the effects

Table 3
Spearman's rank correlation coefficients (rho). Strong positive correlations (> 0.7) are shown in red and weak negative ones (-0.3 to -0.5) in blue. The three most positively correlated indicators with respect to the indicator set are highlighted in orange and the three most negatively correlated ones in green.

Spearman's rho		1 - SAIDI	2 - Severe accident risks	3 - Control of corruption	4 - Political stability and absence of violence/ terrorism	5 - Electricity mix diversity	6 - Electricity import dependence	7 - Equivalent availability factor	8 - GDP per capita	9- Insurance penetration	10 - Government effectiveness	11 - Average outage time	12 - Ease of doing business	Average correlation coefficients	Number of correlations between -
1 - SAIDI	Correlation Coefficient	1.000	.541**	.616**	.513**	.282**	-.374**	0.129	.690**	.437**	.697**	.492**	.676**	0.427	4
	Sig. (2-tailed)		0.000	0.000	0.000	0.001	0.000	0.130	0.000	0.000	0.000	0.000	0.000		
2 - Severe accident risks	Correlation Coefficient	.541**	1.000	.633**	.511**	.361**	-.329**	.360**	.718**	.420**	.686**	.295**	.654**	0.441	3
	Sig. (2-tailed)	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		
3 - Control of corruption	Correlation Coefficient	.616**	.633**	1.000	.800**	.313**	-.460**	-0.014	.808**	.589**	.940**	.239**	.786**	0.477	4
	Sig. (2-tailed)	0.000	0.000		0.000	0.000	0.000	0.865	0.000	0.000	0.000	0.004	0.000		
4 - Political stability and absence of	Correlation Coefficient	.513**	.511**	.800**	1.000	.234**	-.466**	-0.076	.721**	.484**	.778**	.180*	.685**	0.397	5
	Sig. (2-tailed)	0.000	0.000	0.000		0.005	0.000	0.374	0.000	0.000	0.000	0.033	0.000		
5 - Electricity mix diversity	Correlation Coefficient	.282**	.361**	.313**	.234**	1.000	-.188*	-.195*	.356**	.264**	.393**	.211*	.415**	0.222	8
	Sig. (2-tailed)	0.001	0.000	0.000	0.005		0.026	0.021	0.000	0.002	0.000	0.012	0.000		
6 - Electricity import dependence	Correlation Coefficient	-.374**	-.329**	-.460**	-.466**	-.188*	1.000	-0.085	-.389**	-.232**	-.487**	-0.149	-.413**	-0.325	6
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.026		0.319	0.000	0.006	0.000	0.078	0.000		
7 - Equivalent availability factor	Correlation Coefficient	0.129	.360**	-0.014	-0.076	-.195*	-0.085	1.000	0.104	-0.048	0.054	0.001	0.001	0.021	12
	Sig. (2-tailed)	0.130	0.000	0.865	0.374	0.021	0.319		0.220	0.575	0.527	0.992	0.988		
8 - GDP per capita	Correlation Coefficient	.690**	.718**	.808**	.721**	.356**	-.389**	0.104	1.000	.658**	.870**	.318**	.781**	0.512	3
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.220		0.000	0.000	0.000	0.000		
9- Insurance penetration	Correlation Coefficient	.437**	.420**	.589**	.484**	.264**	-.232**	-0.048	.658**	1.000	.626**	.227**	.566**	0.363	6
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.002	0.006	0.575	0.000		0.000	0.007	0.000		
10 - Government effectiveness	Correlation Coefficient	.697**	.686**	.940**	.778**	.393**	-.487**	0.054	.870**	.626**	1.000	.312**	.858**	0.521	3
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.527	0.000	0.000		0.000	0.000		
11 - Average outage time	Correlation Coefficient	.492**	.295**	.239**	.180*	.211*	-0.149	0.001	.318**	.227**	.312**	1.000	.324**	0.223	9
	Sig. (2-tailed)	0.000	0.000	0.004	0.033	0.012	0.078	0.992	0.000	0.007	0.000		0.000		
12 - Ease of doing business	Correlation Coefficient	.676**	.654**	.786**	.685**	.415**	-.413**	0.001	.781**	.566**	.858**	.324**	1.000	0.485	3
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.988	0.000	0.000	0.000	0.000			

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Table 4
List of normalization methods used in this research (Joint Research Centre of the European Commission, 2008; Nelder, 1961).

Normalization method		Formula	Description	Comments
Ordinal	Rank	$I_{ic} = Rank(x_{ic})$	The countries are ranked according to the indicators' values.	Simple and straightforward normalization method. Not affected by outperformers as the scale is changed to ordinal numbers and the distances between performance of countries are lost. Countries performing significantly better than others are disadvantaged.
	Percentile rank	$I_{ic} = \dfrac{a_i + 0.5 f_{x_{ic}}}{N}$	Represents the percentage of scores that are equal to or lower than itself.	Same comments than for the rank method.
	Categorical (a, b, c) where a, b, c is equal to either (−1, 0, 1) or (0.1, 1, 2)	$I_{ic} = \begin{cases} a & if\, x_{ic} < \bar{x}_{ic=c} - \bar{\sigma}_{ic=c} \\ b & if\, \bar{x}_{ic=c} - \bar{\sigma}_{ic=c} \le x_{ic} \le \bar{x}_{ic=c} + \bar{\sigma}_{ic=c} \\ c & if\, \bar{x}_{ic=c} + \bar{\sigma}_{ic=c} < x_{ic} \end{cases}$	Transforms the data according to predefined rules. In the present study, the rules depend on the mean and standard deviations of the indicators across all countries.	Categorical scales can be seen as ordinal values where the levels of performance are categorized into several tiers. For example, low (−1), medium (0) and high (1).
	Categorical (0.1, 0.2, 0.4, 0.6, 0.8, 1)	$I_{ic} = \begin{cases} 0.1 & if\, x_{ic} < \bar{x}_{ic=c} - 2\bar{\sigma}_{ic=c} \\ 0.2 & if\, \bar{x}_{ic=c} - 2\bar{\sigma}_{ic=c} \le x_{ic} < \bar{x}_{ic=c} - \bar{\sigma}_{ic=c} \\ 0.4 & if\, \bar{x}_{ic=c} - \bar{\sigma}_{ic=c} \le x_{ic} < \bar{x}_{ic=c} \\ 0.6 & if\, \bar{x}_{ic=c} \le x_{ic} < \bar{x}_{ic=c} + \bar{\sigma}_{ic=c} \\ 0.8 & if\, \bar{x}_{ic=c} + \bar{\sigma}_{ic=c} \le x_{ic} < \bar{x}_{ic=c} + 2\bar{\sigma}_{ic=c} \\ 1.0 & if\, \bar{x}_{ic=c} + 2\bar{\sigma}_{ic=c} \le x_{ic} \end{cases}$		
Linear	Standardization (z-score)	$I_{ic} = \dfrac{x_{ic} - \bar{x}_{ic=c}}{\bar{\sigma}_{ic=c}}$	The standardization is a linear transformation of the data set with a mean of zero and a standard deviation of one.	Preserves the distribution of the values of the indicators. The standardized data set is not bounded (there are no fixed minimums or maximums).
	Min-max	$I_{ic} = \dfrac{x_{ic} - min(x_i)}{max(x_i) - min(x_i)}$	Normalizes the data between zero and one by applying a linear transformation.	Most common normalization method. New data points that are out of the starting scale can affect the scores of all the indicators. In the presence of outperformers, most of the countries are normalized to very closely related values. The ratios are conserved.
	Target	$I_{ic} = \dfrac{x_{ic}}{max(x_i)}$	Normalizes the upper limit of the data to one.	
Sigmoid	Logistic	$I_{ic} = \dfrac{1}{1 + e^{-\frac{x_{ic} - \bar{x}_{ic=c}}{\bar{\sigma}_{ic=c}}}}$	Normalizes the data into a sigmoid curve (S-shaped) between 0 (for −inf) and 1 (for + inf).	This normalization method reduces the effect of outperformers.

x_{ic}: the value of indicator i for country c.
I_{ic}: the normalized value of indicator i for country c.
a_i: the amount of values lower or equal to itself.
$f_{x_{ic}}$: the frequency of countries with same value x_{ic}.
N: the number of alternatives (140).
$\bar{x}_{ic=c}$: the average value of indicator i across all countries.
$\bar{\sigma}_{ic=c}$: the standard deviation of indicator i across all countries.
$min(x_i)$: the minimum value of indicator i across all countries.
$max(x_i)$: the maximum value of indicator i across all countries.

Table 5
List of aggregation functions used in this research (Langhans et al., 2014; De Condorcet, 2014).

Aggregation function	Formula	Level of compensation	Comments
Additive	$score_c = \sum_{i=1}^{n} w_i I_{ic}$	Full	One of the most common aggregation function. Equals the weighted-sum with equal weights. Suitable if the decision-makers' preference values are linear, meaning that the decision-makers accept that the performance of indicators can compensate each other. For example, low-performing indicators can be fully compensated by high-performing indicators.
Geometric	$score_c = \prod_{i=1}^{n} I_{ic}^{w_i}$	Partial	In many real-world cases, full compensation between indicators is not sought, and such non-compensatory preferences of decision-makers are better represented by other functions than the additive one. The geometric function only allows partial compensation. This function can be problematic if normalized indicators' values are 0 (lowest performing indicator), as the function leads to a score of 0. Hence, only used with normalized data sets containing strictly positive numbers.
Harmonic	$score_c = \dfrac{1}{\sum_{i=1}^{n} \frac{w_i}{I_{ic}}}$	Partial (less than geometric)	This function further reduces the compensation levels allowed compared to the geometric one. The harmonic function cannot be used if normalized indicator's values are 0 (divisions by 0 are undefined). Hence, only used with normalized data sets containing strictly positive numbers.
Minimum	$score_c = \min(I_{1c}, I_{2c}, \ldots, I_{nc})$	None	The lowest performing indicator equals the final score. Particularly suitable if stakeholders want the assessment to be driven by the worst performing indicator. No compensation level is allowed at all.
Median	$score_c = I(I_{1c}, I_{2c}, \cdots, I_{nc})$	Depends on the distribution of the indicators' values	Allows to identify overall trends as one half of a country's indicators are above and the other half below the median. If the performances of the indicators are inhomogeneous, the median function can result in very different rankings compared to the additive function. Low-performing indicators can be overcompensated by well-performing indicators, if the number of well-performing indicators is larger than the low-performing ones. In this case, the median function allows for even more compensation than the additive function.
Condorcet methods	Pairwise comparison between countries	Depends on the distribution of the indicators' values	No data normalization required and no explicit indicator aggregation performed. Comparison based on: 1 The amount of country duels won: For example, each indicator of country A is confronted with each indicator of country B. If the performance of i_1 for country A is better than for country B, country A gets one point. This process is then repeated for all indicators. If country A receives more points than B, it won the country duel. A is then confronted in the same way to all other countries and the amount of duels won (as a percentage) can be calculated. 2 The amount of votes the country has got summed up over all the indicators in all country duels: For example, each indicator of country A is confronted with the ones from country B and the percentage of better performing indicators is calculated. Country A is then confronted with all the other countries in the same procedure. In the end, the average of the better performing indicators in all country duels is taken for the ranking purpose.

score_c: composite resilience score for country c.
n: the number of indicators (12).
w_i: the weight of indicator i.
min(): minimum value of all the indicators.
I: median of the indicators' values.

11

Table 6
List of the 38 combinations of normalization methods and aggregation functions used to develop the ESRI.

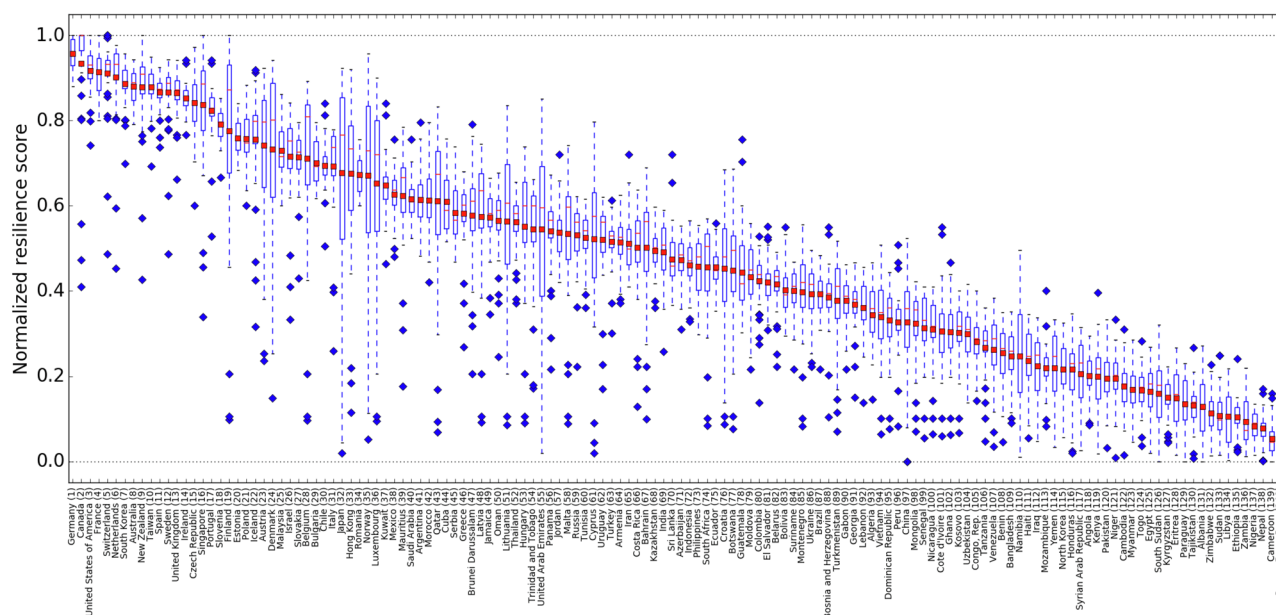| Aggregation function	Normalization method	Comments
Additive	Rank Percentile rank Standardized Min-max Target Logistic Categorical (−1, 0, 1) Categorical (0.1, 0.2, 0.4, 0.6, 0.8, 1)	Being the most popular function, the additive function was used with all types of normalization methods in order to analyze the widest possible sets of combinations. This allows for a comprehensive analysis of the effect of different normalization methods on the final rankings.
Geometric	Percentile rank Standardized + 3 Min-max 0.1–1 Min-max 0.01–1 Target 0.1 Logistic Categorical (0.1, 1, 2) Categorical (0.1, 0.2, 0.4, 0.6, 0.8, 1)	The geometric function was also used with all types of normalization methods, except the rank normalization because multiplying ranks does not give any additional information compared to the combination Geometric – Percentile rank (the final scores are almost the same). As the geometric function does not cope with negative or null values, the standardized data set was linearly shifted to positive numbers by adding 3 ("Standardized + 3") and the ternary categorical scale changed to (0.1, 1, 2). Also, the min–max normalization method was adapted to the ranges [0.1–1] and [0.01–1] instead of [0–1]. This allows analyzing the effects of different compensation levels (less compensation for the range [0.01–1] instead of [0.1–1]). Finally, the target method was linearly shifted to set the minimum value to 0.1 in order to exclude potential zeros.
Harmonic	Percentile rank Standardized + 3 Min-max 0.1–1 Min-max 0.01–1 Target 0.1 Logistic Categorical (0.1, 1, 2) Categorical (0.1, 0.2, 0.4, 0.6, 0.8, 1)	The harmonic function was applied with the same normalization methods than the geometric function due to the same reasons. It allows slightly less compensation than the geometric function.
Minimum	Standardized Logistic	The minimum function was applied to the standardized data set, as the normalized indicators have an equal mean and standard deviation, bringing them to the same scale and because the normalized data is not bounded. Furthermore, the logistic normalization method was also applied as the extreme value of 0 is not included in the normalized data set. The rank, percentile rank, min–max, target and categorical normalization methods bring many countries to the same minimum values, therefore leading to uninterpretable and potentially misleading results. In fact, equally normalized minimum values for the performances of the indicators result in many countries having the same final ranks. Therefore, as the standardized and logistic data set have no equal minimum value, they are particularly suitable for the minimum aggregation function.
Median	Percentile rank Standardized Min-max Target Logistic	The median function was only applied to the percentile rank, standardized, min-max, target and logistic methods. In fact, the rank method is similar to the percentile rank one and due to the large amount of equal values, the categorical scales are not suitable.
Additive/minimum Additive/geometric Additive/median Geometric/harmonic Geometric/minimum Condorcet – Most duels won Condorcet – Most votes	Standardized Standardized + 3 Standardized Standardized + 3 Standardized + 3	According to the decision-maker's preferences, one may want to combine different functions simultaneously. For example, if a decision-maker does not want to have full compensation, but neither the downsides of the less compensatory approaches, he could consider mixes of aggregation functions such as additive-geometric or additive-minimum. The standardized data set was used for these mixes, as it is the one used over all aggregation functions.

Fig. 3. Boxplot of normalized resilience scores per country (data points). For each country, the red square is the value of the Electricity Supply Resilience Index (ESRI). The red horizontal line is the median over the data points. A box contains 50% of all data points, that is, it vertically extends from the first quartile to the third quartile of all data points. The whiskers extend to 3/2 of the length of the box or to the outermost data point, whichever is smaller. The blue rhombi are out-performers.

of normalization methods and aggregation functions on the rankings. Hence, 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 weights. Equal weights also represent the most common profile for such a comparison (Ang et al., 2015a; El Gibari et al., 2018). However, by selecting equal weights with different normalization methods, it is accepted that the trade-offs between the indicators (also called marginal rates of substitution) are not conserved (Gasser et al., 2019). Consecutively, the aggregation combines the normalized indicators data with their respective weights into an index. Many normalization methods and aggregation functions are reported in the literature (Joint Research Centre of the European Commission, 2008). Therefore, rankings of indices are dependent on which combination of normalization and aggregation is used (Wulf et al., 2017). In order to assess the robustness of the measure of electricity supply security resilience of each country, multiple rankings were derived by combining eight normalization methods (rank, percentile rank, standardization, min–max, target, logistic, categorical ternary and categorical senary; see Table 4) and six aggregation functions (additive, geometric, harmonic, minimum, median and Condorcet; see Table 5). The normalization methods denote different preferences from decision-makers, from exploiting only the ordinal character of the data to considering quantitative differences between performances (cardinal character). Furthermore, the aggregation functions represent different levels of compensation between indicators, i.e. if a decision-maker is willing to allow that a low performance of one indicator can be compensated by another indicator fully, to a certain extent or not at all.

Following this methodology, ranking differences are likely to arise. In fact, as the scale of each indicator varies, the normalization methods, excluding the ordinal ones, also hold different levels of compensation between indicators[4]. By using a single weighting profile, this potentially

results in ranking differences. It is possible to keep the rankings equal by adjusting the weighting profiles for each combination so that the compensation levels between indicators remain the same (Gasser et al., 2019; Billaut et al., 2009). However, this was not the goal of the present study, but represents material for further research.
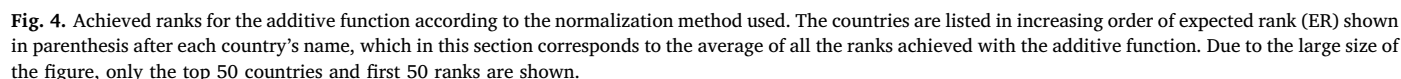
2.5.1. Combinations considered

Indices are constructed through the combination of a normalization method and an aggregation function. In order to study the stability of the scores and resulting ranks, this study uses a wide set of combinations to construct the index so that the robustness of the results can be assessed as well as the variability of the rankings (step 13 in Fig. 1). Due to the nature of the normalization methods and aggregation functions, not all combinations do make sense and some combinations are redundant. The final 38 combinations retained in this study are shown in Table 6, with details about why each combination is relevant and why some combinations were not considered.

2.6. Comparison of scores and rankings

Once the results were calculated for the 38 combinations, a comparison of the scores and ranks was conducted (step 14 in Fig. 1). For this purpose, three ranking comparison measures were used. First, the variability of the scores was assessed through their distribution and visualized using a boxplot. This allows identifying general trends and country groupings. Second, the RDM analyzes a pair of rankings from the perspective of the ranks (actual positions). It compares a sum of differences between the ranks attained by all alternatives in the two rankings with a maximal possible difference of ranks for a pair of orders involving a given number of alternatives. RDM takes values between 0 and 100%, where 100% means that the observed differences are the greatest possible (equivalent to a completely reversed ranking), and 0% indicates that all alternatives attain the same ranks in the compared orders. For example, when RDM is equal to 10%, it means that the sum of rank differences for all alternatives is equal to 10% of the maximal possible differences being observed when one ranking negates the other. For a formal definition of RDM, see Kadziński and Michalski (2016). Third, the RAIs, i.e. the likelihood that a country ranks at a

---

[4] The levels of compensation between indicators represent the necessary increase in performance of an indicator to balance the decrease of another indicator. In other words, it is the Marginal Rate of Substitution (MRS), i.e. the amount of increase needed in an indicator to compensate a unitary decrease in another indicator, in order to keep the same performance on the index.

Fig. 4. Achieved ranks for the additive function according to the normalization method used. The countries are listed in increasing order of expected rank (ER) shown in parenthesis after each country's name, which in this section corresponds to the average of all the ranks achieved with the additive function. Due to the large size of the figure, only the top 50 countries and first 50 ranks are shown.

Table 7

RDM for the eight normalization methods taken with the additive aggregation function, with their mean and standard deviation (SD) indicated. The RDM for the 38 combinations is given in Table S11.

Same aggregation (additive) and different normalization									
RDM	**Additive - Rank**	**Additive - Percentile rank**	**Additive - Standardized**	**Additive - Min-max**	**Additive - Target**	**Additive - Logistic**	**Additive - Categorical -1, 0, 1**	**Additive - Categorical senary**	**Mean ± SD**
Additive - Rank	0.08%	1.96%	5.81%	6.39%	7.22%	5.35%	16.26%	8.83%	6.49 ±4.85%
Additive - Percentile rank	1.96%	0.02%	5.94%	6.45%	7.15%	5.35%	16.33%	8.64%	6.48 ±4.87%
Additive - Standardized	5.81%	5.94%	0.00%	2.14%	4.53%	2.04%	13.39%	6.69%	5.07 ±4.09%
Additive - Min-max	6.39%	6.45%	2.14%	0.00%	3.69%	2.63%	13.54%	6.83%	5.21 ±5.21%
Additive - Target	7.22%	7.15%	4.53%	3.69%	0.00%	4.69%	14.28%	8.16%	6.22 ±4.15%
Additive - Logistic	5.35%	5.35%	2.04%	2.63%	4.69%	0.00%	13.97%	6.79%	5.10 ±4.20%
Additive - Categorical -1, 0, 1	16.26%	16.33%	13.39%	13.54%	14.28%	13.97%	5.78%	11.91%	13.18 ±3.33%
Additive - Categorical senary	8.83%	8.64%	6.69%	6.83%	8.16%	6.79%	11.91%	2.50%	7.54 ±2.66%

certain position, were computed. RAIs are presented in Section S11 and Fig. S1.

The rank comparison measures of normalized hit ratio, Kendall's tau, rank agreement measure and pairwise winning index were also computed (see Kadziński and Michalski (2016) for a description of these), but are not shown in this paper. In fact, these additional measures are not as descriptive as the three hereby considered, hence they do not lead to additional insights.

3. Results, discussion and policy implications

For each country, 38 combinations of normalization methods and aggregation functions result in specific electricity supply resilience scores. Fig. 3 shows the scores of the 140 countries, normalized with the min–max method (the data for the raw scores and normalized scores are given in Tables S8 and S9). The countries are classified in descending order according to the average over the 38 min–max normalized combinations, which is hereby defined as the Electricity Supply Resilience Index (ESRI). Using ESRI, an electricity supply resilience trend can be clearly identified. This proves that ESRI is an accurate and
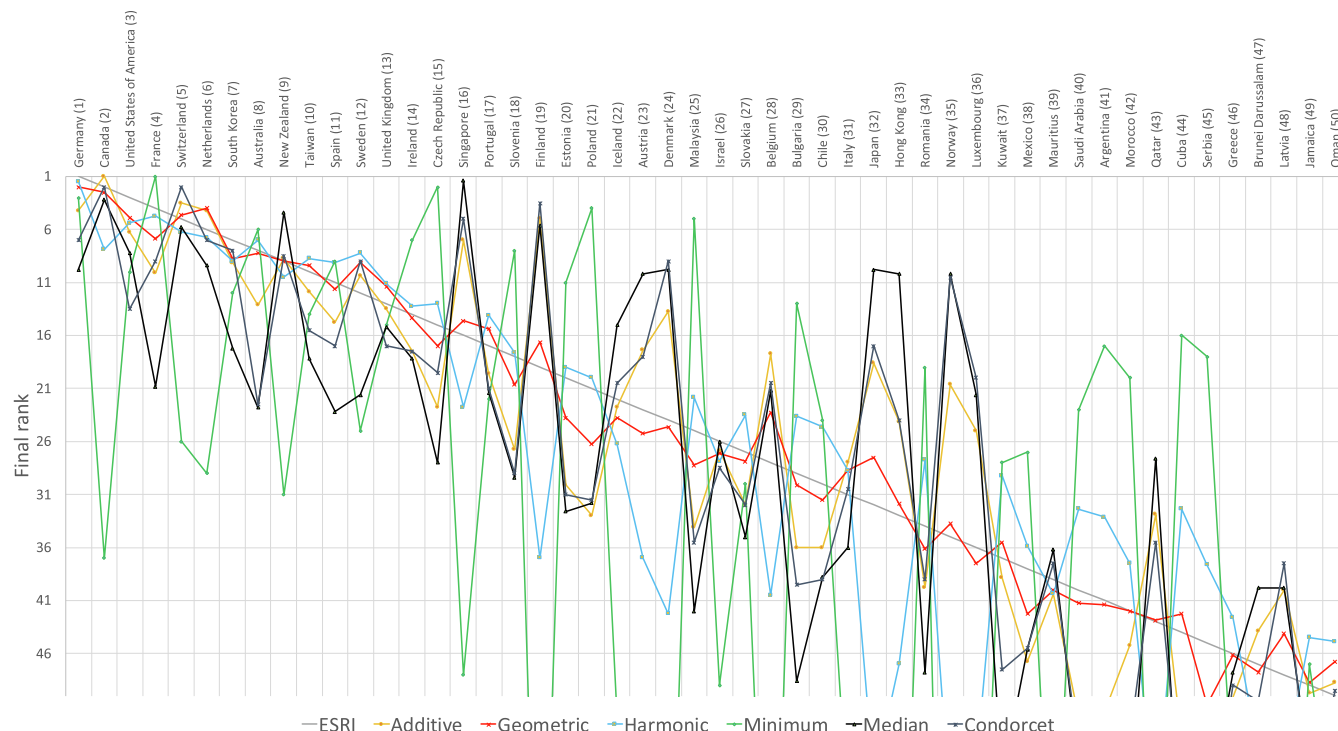
Fig. 5. Achieved ranks according to the aggregation function used (averaged over all the normalization methods considering it). The countries are listed in increasing order of ESRI rank shown in parenthesis after each country's name. Due to the large size of the figure, only the top 50 countries and first 50 ranks are shown.

appropriate measure.

Furthermore, distinct country groupings can be observed, with several countries scoring at the top or at the bottom across all combinations. On the one end, Germany (1[5]), which scores at least 0.88 in all combinations, is the top performer. Next to Germany (1), the following countries to score at least 0.9 on average are Canada (2), the United States of America (USA) (3), France (4), Switzerland (5) and the Netherlands (6). All these countries have top performing indicators and even their worst ones still outperform those from many other countries. On the other end, the Democratic Republic of Congo (DRC) (140), Cameroon (139), Nepal (138) and Nigeria (137) all have scores lower than 0.2 for all combinations and their ESRI is lower than 0.1. None of these countries have enough well-performing indicators 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 indicators do not have much influence on the final scores.

However, the scores tend to be more variable for average-performing countries or countries with divergent indicator performances. The large disparity of the scores proves that the final rankings can have a strong dependency on the considered normalization/aggregation combination. The disparity is especially large for Japan (32), Norway (35) and Finland (19), for which the score differences go up to 0.9. This is due to the fact that each of these countries has simultaneously top- and bottom-performing indicators. Hence, this divergence of performances emphasizes the different levels of compensation allowed by the aggregation functions, leading to large score disparities. On the opposite side, Germany (1), Romania (34), and the DRC (140), for example, show narrow ranges of scores. These patterns are further discussed in the two subsequent sections.

3.1. Effects of normalization methods on the final ranking

The effect of different normalization methods can be studied by comparing the indices and rankings obtained with the same aggregation function, but different normalization methods. The aim is to allow decision-makers to assess the implications of using different strategies to compare the input data, with respect to its ordinal and cardinal characters. Fig. 4 shows the achieved ranks for the additive function according to the eight normalization methods (see Table 4). Overall, the rankings with different normalization methods tend towards similar results, and a trend can thus be identified. The biggest difference emerges with categorical ternary. In fact, as the categorical ternary normalized data set only contains three different values (−1, 0 and 1), there are many ties between the countries. Therefore, this normalization method only poorly differentiates the performances of the indicators and consequently leads to poorly differentiated scores and ranks. Such a normalization method is more suitable when the exact performance on the indicators is unknown and where the best available option is to set ordinal scales (e.g. high, medium and low performance).

In addition, Table 7 illustrates the RDM for the additive function only. The rank and percentile rank normalization methods are the closest as they show an RDM of 1.96%. This is due to the fact that both methods rely on ranks, and the percentile rank only adds a frequency dimension on top. The ranking differences become larger with the standardized, min–max, target and logistic methods, and maximum with the categorical ones. As shown earlier, the biggest difference emerges with categorical ternary (i.e. highest RDM values, with an average of 14.24%), followed by categorical senary (average RDM of 8.26%). Nevertheless, all the numbers in Table 7, hence for the additive aggregation function, are still close to 0, with an average of 6.91% and a standard deviation of maximum 4.87%. Furthermore, the trend is the same for the geometric and harmonic aggregation functions (see Table S11). Finally, considering all aggregation functions, thus the 38 combinations, the overall average RDM is 15.5% with a maximum standard deviation of 9.54%. This confirms that the rankings of all of these combinations are close and, importantly, that a common trend exists.

---

[5] The number inside the parenthesis that follows a country's name indicates its ESRI rank.

Table 8
RDM with the same aggregation functions and same normalization methods.

RDM	Same aggregation				Same normalization		
	Additive - Standardized	Additive - Min-max	Additive - Target	Additive - Logistic	Geometric - Standardized	Harmonic - Standardized	Minimum - Standardized
Additive - Standardized		2.14%	4.53%	2.04%	5.08%	12.12%	35.71%

RDM							
	Additive - Standardized	Additive - Min-max	Additive - Target	Additive - Logistic	Geometric - Logistic	Harmonic - Logistic	Minimum - Logistic
Additive - Logistic	2.04%	2.63%	4.69%		6.00%	14.33%	36.62%

Hence, ESRI is a robust measure to quantify a country's electricity supply resilience.

3.2. Effects of aggregation functions on the final ranking

The effect of different aggregation functions can be studied by comparing the indices and rankings obtained with the same normalization method, but different aggregation functions. Fig. 5 shows the achieved ranks according to the aggregation function used (averaged over all the normalization methods considering them). Even though the rankings are more variable compared to the effects of normalization methods, a trend can still be identified, from top-performing to low-performing countries. The variability seems to increase, especially for average-performing countries or countries with divergent indicator performances. The biggest differences emerge with the minimum function, followed by the median and harmonic functions.

Countries that show rather large rank disparities are the most appropriate ones to analyze the effects of aggregation functions. For example, the United Arab Emirates' (UAE) rank (55) is better for the additive aggregation function (average rank of 36 across the eight additive combinations), because it allows for full compensation between indicators (see Table 6 and Table S10). Its average rank for the geometric aggregation is 58, which shows that this function allows for less compensation than the additive one. Regarding the harmonic function, it allows even less compensation than the geometric one (Langhans et al., 2014), resulting in an average rank of 87 over the eight harmonic combinations. Finally, the minimum function is the one that allows the least compensation and results in an average rank of 124. The best ranks for the UAE are achieved through the median function (average rank of 24). In fact, as more than half of its indicators rank almost at the top, the performance of the rest of the indicators becomes merely irrelevant. Therefore, in a case where most of the indicators are well-performing, the median function over-compensates the low-performing indicators, in the sense that the ranks are even higher than for the additive function, which already offers full compensation. Regarding improvement possibilities, the UAE's low-performing indicators are the electricity mix diversity (i_5) and the average outage time (i_{11}). In fact, more than 98% of its electricity is produced by natural gas and each interruption lasts 8 hours on average. This could be improved by, e.g. promoting renewable energies and developing state-of-the-art emergency preparedness measures and protocols to apply in case of disruptions. Japan (32), Norway (35) and Singapore (16), among others, show the same rank pattern, also due to their divergent indicator performances.

One particularity of the minimum function is that it was combined with the standardized and logistic normalization methods only. In fact, the minimum function should be applied with data sets that have no lower bound. Therefore, it results in different values for each country, which allows ranking them without having equally ranked countries.

Furthermore, these combinations strongly penalize countries that have very low performance on a single indicator compared to the other countries. In fact, if a country has such performance, its standardized or logistics values will be very low too[6]. This is the case for Albania (131) and China (97), where for severe accidents there are 7.03 fatalities per GWeyr (i_2). This value is by far worse compared to the other countries and the variance in terms of performance for this indicator is larger than the variance of the other indicators (the other indicators do not show such high indicator performance divergence).

The decreasing levels of compensation allowed by the additive, geometric, harmonic and minimum functions are also shown in the RDMs (see Table 8 and Table S11). Using the standardized data set for comparison, as it is considered over all these aggregation functions, the RDM between the additive and geometric functions is 5.08%, between the additive and harmonic 12.12% and between the additive and minimum 35.71%. As these values are increasing, they show an increasing discrepancy between the rankings, which confirms the gradually decreasing levels of compensation. The same pattern is found for the logistic normalization method. Furthermore, the RDM values when comparing different aggregation functions are lower than the ones obtained when comparing different normalization methods. Therefore, it can once again be concluded that the aggregation functions have a higher influence on the variability of the results.

Finally, the two rankings constructed with the Condorcet methods (combinations 37 and 38) can provide additional insights as in the cases of Japan (32), Norway (35) and the UAE (55). Due to some low-performing indicators (i_6 and i_{11} for Japan, i_5 and i_7 for Norway and i_5 and i_{11} for the UAE), these three countries do not rank at the highest level when counting the number of better performing indicators (equivalent to the number of votes, see combination number 38). However, the number of bilateral country duels won is still high (a country duel is won when more than half of the indicators of a country are better than the ones from another country), because most of their indicators are still top-performing (see combination number 37). These results are supported by the fact that these countries show among the highest standard deviations for the normalized indicator scores, indicating that the performances of their indicators are not evenly distributed (i.e. these countries score at the top for some indicators, but also at the bottom for other ones).

Compared to the countries mentioned above, an opposite trend is found for Romania (34), Estonia (20) and Ecuador (75), which have most indicators performing at an average level. Such performances do not result in a high rank for the additive, median and Condorcet functions. Nevertheless, these countries are not comparatively penalized as much by the geometric, harmonic or minimum functions,

---

[6] On the other hand, if the performance of all the countries is close, the minimum and logistic normalized values comparatively will not be as low.

because they do not have as many low-performing indicators as other countries (including Japan (32), Norway (35) and the UAE (55)). Therefore, countries without large disparities in the indicator performances show smaller rank disparities and gradually better ranks for the geometric, harmonic and minimum functions.

Combinations 32 to 36 involve a uniformly weighted mix of two aggregation functions. Such combinations can be applied if a certain degree of compensation is desired. For example, combination number 32 provides a 0.5 compensation level, as the additive function allows full compensation and the minimum one none. The weights of the two aggregation functions could be varied to allow any level of compensation between 0 and 1. As expected, the final rankings of these combinations are situated in between the rankings achieved by considering the aggregation functions separately.

The choice of the aggregation function can be related to the decision-makers' preferences. For example, if the decision-makers agree to allow full compensation between performances, they can opt for the additive aggregation function. However, the resilience of systems might be more related to their weak components, since a failure in one area might be severe enough to spread through an entire network. If the decision-makers advocate for the latter system representation, they should opt for a geometric or harmonic aggregation function that allows less compensation. In extreme cases, under the motto "as strong as the weakest link", the minimum function should be considered as it emphasizes the single worst-performing indicator.

4. Conclusions

The objectives of this study were threefold. First, an indicator set that comprehensively represents the electricity supply resilience of countries, i.e. the four dimensions, was developed for the first time. It combines technical indicators with measures related to governance, geopolitical and organizational factors. Second, the statistical coherence of the indicator set was verified through correlation analysis and reliability assessment, confirming that indices can be constructed from the available data set. Third, these indices were developed by considering 38 combinations of eight normalization methods and six aggregation functions, in order to assess the robustness of the results. The normalization methods denote different preferences from decision-makers, from exploiting only the ordinal character of the data to considering quantitative differences between performances (cardinal character). Furthermore, the aggregation functions represent different levels of compensation between indicators. Finally, the Electricity Supply Resilience Index (ESRI) is proposed and quantified for 140 countries worldwide. The ESRI is the first comprehensive country performance measure for electricity supply resilience, which allows to build a univocal country ranking.

A clear electricity supply security resilience ranking trend across all the investigated combinations is apparent. The ESRI robustly ranks the same countries at the top (e.g. Germany, Canada and the USA), whereas others constantly rank at the bottom (e.g. the DRC, Cameroon, Nepal and Nigeria). In contrast, the rankings are more variable for average-performing countries or countries with divergent indicator performance. Consequently, the ESRI helps decision-makers to identify trends and specific patterns, and to understand the responsible factors behind.

The novel robustness assessment demonstrated that the ranks of the investigated countries are primarily affected by the aggregation function used, and to a lesser extent by the normalization method. This is confirmed by the Rank Difference Measure (RDM), which has never been used before in the energy sector. The disparity is particularly large for countries having a few top- and bottom-performing indicators, which is explained by the variable levels of compensation allowed by the different aggregation functions. Hence, countries with (even only one) low-performing indicators score better if approaches allowing for higher or full compensation are applied. Overall, the RDM has been proven to be a useful tool to study ranking robustness and it confirmed

that countries on average maintain their ESRI rank. Furthermore, it identified the categorical normalization method as the most deficient one, due to its poor capacity to differentiate performances. These findings are confirmed with the untrimmed data set, therefore proving that data trimming did not affect the results.

This research demonstrates the dependence and sensitivity of the results on methodological choices in the construction of indices. By exploring the robustness of the indices from a wider perspective, it challenges the status quo of developing indices based on a single combination, to accommodate a variety of approaches, even extreme ones, and proves that for this case study the results are robust. In fact, decision-makers need to be aware of the characteristics of the different approaches and how they can potentially affect the results. Selecting an approach implicitly defines the indicator compensation levels, which, in turn, affects the index and the derived rankings and conclusions. Therefore, the novel robustness analysis presented herein greatly improves the credibility of the results.

Another contribution of this paper is that it allows decision-makers to understand on which indicators they should focus to improve the performance of their country, and therefore achieve a better overall ranking. If the ranking is constructed with the additive function, an improvement of any indicator will have the same effect on the final ranks, as the additive function is linear and allows full compensation. However, using the geometric, harmonic or minimum functions implies that an improvement of a low-performing indicator will have a larger impact on the final ranks compared to the situation where an already well-performing indicator is improved by the same amount. Of course, this needs to be complemented by a reality check of which indicators a country can and wants to improve, and what this means in terms of costs, feasibility and political willingness.

Overall, the use of an appropriate index enables decision-makers to get a clear understanding of the problem to be studied. Furthermore, it is important for them to analyze the consequences and resulting biases of the chosen index' construction methodology and a robustness analysis should be performed. Indices are a particularly promising approach for electricity supply resilience quantification, because of their ability to assess the multi-dimensional aspects of resilience. Finally, the present case study confirms that the ESRI constitutes an appropriate, accurate, consistent and robust measure to quantify countries' electricity supply resilience.

Future research could address the following topics. First, a potential expansion of the indicator data set to include and study the impact of currently omitted indicators. Examples include, among others, price volatility, reserves to consumption ratio or the technical qualified personnel available in the energy sector. Second, instead of only considering yearly averages, the effects on the resilience of electricity supply of the different seasonal consumption profiles could be studied. Third, extending the scope of the present study, the long-term sustainability of a country's electricity mix could be determined. In fact, countries can have similar diversity indicator values, but the sustainability of the technologies used may vary greatly. Fourth, instead of considering equal weights, the implicit weights due to the correlations between indicators could be studied (Becker et al., 2017; Maxim, 2014), along with the weighting profiles required to get a single ranking over all combinations (Billaut et al., 2009). Fifth, stakeholders (e.g. plant operators, governmental organizations, etc.) could be involved to determine the indicator weighting profiles that suits them best. Sixth, it would be interesting to compare the present ranking with existing ones, not only with respect to energy security, but also in different fields (e.g. from the World Development Indicators (WDI) database (World Bank, 2017)), to potentially identify similarities and correlations among different country rankings.

Becker, Dr. Marcos Dominguez-Torreiro and Dr. Michaela Saisana from the Joint Research Centre of the European Commission for their insightful feedback provided during the conduct of this research project.

The research was conducted at the Future Resilient Systems (FRS) at the Singapore-ETH Centre (SEC), which was established collaboratively between ETH Zürich and Singapore's National Research Foundation (FI 370074011) under its Campus for Research Excellence And Technological Enterprise (CREATE) program. Dr. Miłosz Kadziński acknowledges support from the Polish Ministry of Science and Higher Education under grant no. 0296/IP2/2016/74.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2019.105731.

References

Ang, B.W., Choong, W.L., Ng, T.S., 2015a. Energy security: definitions, dimensions and indexes. Renew. Sustain. Energy Rev. 42, 1077–1093.
Ang, B.W., Choong, W.L., Ng, T.S., 2015b. A framework for evaluating Singapore's energy security. Appl. Energy 148, 314–325.
Angelis-Dimakis, A., Arampatzis, G., Assimacopoulos, D., 2012. Monitoring the sustainability of the Greek energy system. Energy Sustain. Dev. 16 (1), 51–56.
Antanasijević, D., Pocajt, V., Ristić, M., Perić-Grujić, A., 2017. A differential multi-criteria analysis for the assessment of sustainability performance of European countries: beyond country ranking. J. Cleaner Prod. 165, 213–220.
Asgary, A., Ozdemir, A. I., Gentles, C., 2015. Does Insurance Delay or Speed up the Recovery and Reconstruction Process? Evidences from Canada. In: Reconstruction and Recovery in Urban Contexts. London, United Kingdom.
Asia Pacific Energy Research Centre (APERC), 2007. Quest for Energy Security in the 21st Century—Resources and Constraints. Asia Pacific Energy Research Centre, Institute of Energy Economics, Tokyo, Japan.
Asian Development Bank. Key Indicators for Asia and Pacific. 2017 [retrieved on 3 January 2018]; Available from: https://www.adb.org/data/statistics.
Augutis, J., Krikštolaitis, R., Matuziene, V., Pečiulytė, S., 2009. Assessment of Lithuanian power supply security depending on nuclear energy, in Safety and Security Engineering III. pp. 235–247.
Augutis, J., Krikštolaitis, R., Pečiulytė, S., Konstantinavičiūtė, I., 2011. Sustainable development and energy security level after Ignalina NPP shutdown. Technol. Econ. Dev. Econ. 17 (1), 5–21.
Augutis, J., Krikstolaitis, R., Martisauskas, L., Peciulyte, S., 2012. Energy security level assessment technology. Appl. Energy 97, 143–149.
Azzuni, A., Breyer, C., 2017. Definitions and dimensions of energy security: a literature review. Wiley Interdisciplinary Reviews: Energy and Environment.
Badea, A.C., Rocco, S.C.M., Tarantola, S., Bolado, R., 2011. Composite indicators for security of energy supply using ordered weighted averaging. Reliab. Eng. Syst. Saf. 96 (6), 651–662.
Bandura, R., 2008. A Survey of Composite Indices Measuring Country Performance: 2008 Update. United Nations Development Programme, Office of Development Studies (UNDP/ODS Working Paper), New York.
Becker, W., Saisana, M., Paruolo, P., Vandecasteele, I., 2017. Weights and importance in composite indicators: closing the gap. Ecol. Ind. 80, 12–22.
Belton, V., Stewart, T.J., 2002. Multiple Criteria Decision Analysis: An Integrated Approach. Springer Science + Business Media, Dordrecht, Netherlands.
Bergström, J., van Winsen, R., Henriqson, E., 2015. On the rationale of resilience in the domain of safety: a literature review. Reliab. Eng. Syst. Saf. 141, 131–141.
Billaut, J.-C., Bouyssou, D., Vincke, P., 2009. Should you believe in the Shanghai ranking? Scientometrics 84 (1), 237–263.
Blyth, W., Lefevre, N., 2004. Energy Security and Climate Change Policy Interactions: An Assessment Framework. IEA.
Boccauthor, R., Hanna, A., 2016. Global Energy Architecture Performance Index Report. In: World Economic Forum.
Böhringer, C., Jochem, P.E., 2007. Measuring the immeasurable—A survey of sustainability indices. Ecol. Econ. 63 (1), 1–8.
Bompard, E., Carpignano, A., Erriquez, M., Grosso, D., Pession, M., Profumo, F., 2017. National energy security assessment in a geopolitical perspective. Energy.
BP, 2017. Review of World Energy, Petroleum, B., Editor.
British Petroleum, BP Statistical Review of World Energy June 2017, 2017, http://www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/statistical-review-2017/bp-statistical-review-of-world-energy-2017-full-report.pdf.
Brown, A., Matlock, M.D., 2011. A review of water scarcity indices and methodologies. White paper 106, 19.
Brown, M.A., Wang, Y., Sovacool, B.K., D'Agostino, A.L., 2014. Forty years of energy security trends: a comparative assessment of 22 industrialized countries. Energy Res. Social Sci. 4, 64–77.
Burgass, M.J., Halpern, B.S., Nicholson, E., Milner-Gulland, E.J., 2017. Navigating uncertainty in environmental composite indicators. Ecol. Ind. 75, 268–278.
Burgherr, P., Eckle, P., Hirschberg, S., Cazzoli, E., 2011. Final report on severe accident risks including key indicators. SECURE Deliverable D5, p. 7.2.

Burgherr, P., Hirschberg, S., 2008. A comparative analysis of accident risks in fossil, hydro, and nuclear energy chains. Hum. Ecol. Risk Assess. 14 (5), 947–973.
Burgherr, P., Hirschberg, S., 2014. Comparative risk assessment of severe accidents in the energy sector. Energy Policy 74 (Suppl. 1), S45–S56.
Burgherr, P., Eckle, P., Hirschberg, S., 2012. Comparative assessment of severe accident risks in the coal, oil and natural gas chains. Reliab. Eng. Syst. Saf. 105, 97–103.
Cabalu, H., 2010. Indicators of security of natural gas supply in Asia. Energy Policy 38 (1), 218–225.
Cabalu, H., Alfonso, C., 2013. Energy Security in Asia: The Case of Natural Gas. In: Energy Economics and Financial Markets. pp. 13–30.
Centre for Environmental Law & Policy, Environmental Performance Index, 2018 Technical Appendix, 2018, https://epi.envirocenter.yale.edu/downloads/epi2018technicalappendixv02.pdf.
Cimellaro, G.P., 2016. In: Urban Resilience for Emergency Response and Recovery. Springer, pp. 49–69.
Cinelli, M., 2017. The art of supporting decision-making. Exchanges: Warwick Res. J. 4 (2), 298–308.
Cinelli, M., Coles, S.R., Kirwan, K., 2014. Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment. Ecol. Ind. 46, 138–148.
Cohen, G., Joutz, F., Loungani, P., 2011. Measuring energy security: trends in the diversification of oil and natural gas supplies. Energy policy 39 (9), 4860–4869.
Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. Psychometrika 16 (3), 297–334.
Cutter, S.L., 2015. The landscape of disaster resilience indicators in the USA. Nat. Hazards 80 (2), 741–758.
De Condorcet, N. (Ed.), 2014. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Cambridge University Press, Cambridge.
Dobbie, M.J., Dail, D., 2013. Robustness and sensitivity of weighting and aggregation in constructing composite indices. Ecol. Ind. 29, 270–277.
Dodgson, J., Spackman, M., Pearman, A., Phillips, L., 2009. Multi-Criteria Analysis: A Manual. Department for Communities and Local Government, London.
Doukas, H., Papadopoulou, A., Savvakis, N., Tsoutsos, T., Psarras, J., 2012. Assessing energy sustainability of rural communities using Principal Component Analysis. Renew. Sustain. Energy Rev. 16 (4), 1949–1957.
Dunn, L., Dunn, R., 2012. W&J Energy Index. Washington & Jefferson College.
Eckle, P., Burgherr, P., Hirschberg, S., Final Report on Multi Criteria Decision Analysis (MCDA), 2011, https://www.psi.ch/ta/SecureEN/WP6D2.pdf.
Ediger, V.Ş., Berk, I., 2011. Crude oil import policy of Turkey: historical analysis of determinants and implications since 1968. Energy Policy 39 (4), 2132–2142.
El Gibari, S., Gómez, T., Ruiz, F., 2018. Building composite indicators using multicriteria methods: a review. J. Business Econ.
Erahman, Q.F., Purwanto, W.W., Sudibandriyo, M., Hidayatno, A., 2016. An assessment of Indonesia's energy security index and comparison with seventy countries. Energy 111, 364–376.
European Commission, Ease of Doing Business – Thematic Discussions on Growth and Jobs, 2017, https://www.google.com.sg/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjFlonTmPLVAhXILo8KHSuCCDYQFggmMAA&url=http%3A%2F%2Fwww.consilium.europa.eu%2Fen%2Fmeetings%2Feurogroup%2F2017%2F02%2Fnote-ease-of-doing-business-20170220_pdf%2F&usg=AFQjCNEdBx149XOhup3fD_w4byxNSqLCww.
Field, A., 2013. Discovering Statistics using IBM SPSS Statistics. Sage.
Finster, M., Phillips, J., Wallace, K., 2016. Front-Line Resilience Perspectives: The Electric Grid. http://www.ipd.anl.gov/anlpubs/2017/02/133583.pdf.
Fowler, S.J., Hope, C., 2007. A critical review of sustainable business indices and their impact. J. Bus. Ethics 76 (3), 243–252.
Foxon, T.J., McIlkenny, G., Gilmour, D., Oltean-Dumbrava, C., Souter, N., Ashley, R., Butler, D., Pearson, P., Jowitt, P., Moir, J., 2002. Sustainability criteria for decision support in the UK water industry. J. Environ. Plann. Manage. 45 (2), 285–301.
Francis, R., Bekera, B., 2014. A metric and frameworks for resilience analysis of engineered and infrastructure systems. Reliab. Eng. Syst. Saf. 121, 90–103.
Freudenberg, M., 2003. Composite indicators of country performance.
Frondel, M., Schmidt, C.M., 2008. Measuring energy security-a conceptual note.
Füssel, H.-M., 2010. Review and quantitative analysis of indices of climate change exposure, adaptive capacity, sensitivity, and impacts.
Gan, X., Fernandez, I.C., Guo, J., Wilson, M., Zhao, Y., Zhou, B., Wu, J., 2017. When to use what: methods for weighting and aggregating sustainability indicators. Ecol. Ind. 81, 491–502.
Gasser, P., Cinelli, M., Spada, M., Burgherr, P., Stojadinović, B., 2019. Indices under the spotlight: An approach to unveil and manage the implicit trade-offs between indicators. Eur. Safe. Reliab. Conf In press.
Gasser, P., Lustenberger, P., Sun, T., Kim, W., Spada, M., Burgherr, P., Hirschberg, S., Stojadinović, B., 2017. Security of electricity supply indicators in a resilience context. European Safety and Reliability Conference. Taylor & Francis, Portorož, Slovenia.
Geng, J.-B., Ji, Q., 2014. Multi-perspective analysis of China's energy supply security. Energy 64, 541–550.
Ghasemi, A., Zahediasl, S., 2012. Normality tests for statistical analysis: a guide for non-statisticians. Int. J. Endocrinol. Metab. 10 (2), 486.
Glynn, J., Chiodi, A., Gallachóir., Ó.B., 2017. Energy security assessment methods: quantifying the security co-benefits of decarbonising the Irish Energy System. Energy Strategy Rev. 15, 72–88.
Gnansounou, E., 2008. Assessing the energy vulnerability: case of industrialised countries. Energy Policy 36 (10), 3734–3744.
Greco, S., Ehrgott, M., Figueira, J.R., 2016. Multiple criteria decision analysis: state of the art surveys. Second Edition ed. State of the Art Surveys.[S. l.]: Springer Science & Business Media. Volume 1 and 2. International Series in Operations Research & Management Science. Springer.

Grupp, H., Mogee, M.E., 2004. Indicators for national science and technology policy: how robust are composite indicators? Res. Policy 33 (9), 1373–1384.

Gulati, M., 2006. Corruption in Electricity Sector. http://www1.worldbank.org/publicsector/anticorrupt/GulatiApril6.ppt.

Gupta, E., 2008. Oil vulnerability index of oil-importing countries. Energy Policy 36 (3), 1195–1211.

Häring, I., Sansavini, G., Bellini, E., Martyn, N., Kovalenko, T., Kitsak, M., Vogelbacher, G., Ross, K., Bergerhausen, U., Barker, K., 2017. Towards a generic resilience management, quantification and development process: general definitions, requirements, methods, techniques and measures, and case studies. In: Resilience and Risk. Springer, pp. 21–80.

Hawkins, D.M., 1980. Identification of Outliers 11 Springer.

Heinimann, H.R., Hatfield, K., 2017. Infrastructure resilience assessment, management and governance-state and perspectives. In: Resilience and Risk. Springer, pp. 147–187.

Hirschberg, S., Burgherr, P., Spiekerman, G., Dones, R., 2004. Severe accidents in the energy sector: comparative perspective. J. Hazard. Mater. 111 (1–3), 57–65.

Hirschberg, S., Bauer, C., Burgherr, P., Dones, R., Simons, A., Schenler, W., Bachmann, T., Gallego Carrera, D., 2008. Final set of sustainability criteria and indicators for assessment of electricity supply options. NEEDS deliverable, 2008(D3).

Hosseini, S., Barker, K., 2016. A Bayesian Network Model for Resilience-Based Supplier Selection. Int. J. Prod. Econ.

Hosseini, S., Barker, K., Ramirez-Marquez, J.E., 2016. A review of definitions and measures of system resilience. Reliab. Eng. Syst. Saf. 145, 47–61.

Hu, J.-L., Kao, C.-H., 2007. Efficient energy-saving targets for APEC economies. Energy Policy 35 (1), 373–382.

Huang, I.B., Keisler, J., Linkov, I., 2011. Multi-criteria decision analysis in environmental sciences: ten years of applications and trends. Sci. Total Environ. 409 (19), 3578–3594.

Hughes, L., Shupe, D., 2010. Creating Energy Security Indexes with Decision Matrices and Quantitative Criteria. Energy Research Group, Halifax, Nova Scotia, Canada http://www.worldenergy.org.

Iddrisu, I., Bhattacharyya, S.C., 2015. Sustainable energy development index: a multi-dimensional indicator for measuring sustainable energy development. Renew. Sustain. Energy Rev. 50, 513–530.

Institute for 21st Century Energy, International index of energy security risk, 2016, https://www.globalenergyinstitute.org/sites/default/files/energyrisk_intl_2016.pdf.

Institute for 21st Century Energy, Index of U.S. energy security risk, 2017, https://www.globalenergyinstitute.org/sites/default/files/USIndexFinal2PDF.pdf.

International Energy Agency, Energy supply security 2014, part 1, 2014, https://www.iea.org/publications/freepublications/publication/ENERGYSUPPLYSECURITY2014.pdf.

International Energy Agency. Statistics. 2015 [retrieved on 8 March 2018]; Available from: https://www.iea.org/statistics/statisticssearch.

International Renewable Energy Agency, Renewable Energy Statistics 2017, 2017, https://www.irena.org/publications/2017/Jul/Renewable-Energy-Statistics-2017.

Ishizaka, A., Nemery, P., 2013. Multi-Criteria Decision Analysis: Methods and Software. John Wiley & Sons.

Jasiński, D., Cinelli, M., Dias, L.C., Meredith, J., Kirwan, K., 2018. Assessing supply risks for non-fossil mineral resources via multi-criteria decision analysis. Resour. Policy.

Jewell, J., 2011. The IEA model of short-term energy security (MOSES).

Joint Research Centre of the European Commission, 2008. Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD publishing.

Kadziński, M., Michalski, M., 2016. Scoring procedures for multiple criteria decision aiding with robust and stochastic ordinal regression. Comput. Oper. Res. 71, 54–70.

Kamsamrong, J., Sorapipatana, C., 2014. An assessment of energy security in Thailand's power generation. Sustain. Energy Technol. Assess. 7, 45–54.

Kanchana, K., McLellan, B.C., Unesaki, H., 2016. Energy dependence with an Asian twist? Examining international energy relations in Southeast Asia. Energy Res. Social Sci. 21, 123–140.

Kruyt, B., van Vuuren, D.P., de Vries, H.J.M., Groenenberg, H., 2009. Indicators for energy security. Energy Policy 37 (6), 2166–2181.

Laldjebaev, M., Morreale, S.J., Sovacool, B.K., Kassam, K.-A.S., 2018. Rethinking energy security and services in practice: national vulnerability and three energy pathways in Tajikistan. Energy Policy 114, 39–50.

Langhans, S.D., Reichert, P., Schuwirth, N., 2014. The method matters: a guide for indicator aggregation in ecological assessments. Ecol. Ind. 45, 494–507.

Layton, L., 2015. Electric system reliability indices. Retrieved May, 2004. p. 11.

Le Coq, C., Paltseva, E., 2009. Measuring the security of external energy supply in the European Union. Energy Policy 37 (11), 4474–4481.

Lefèvre, N., 2010. Measuring the energy security implications of fossil fuel resource concentration. Energy Policy 38 (4), 1635–1644.

Li, Y., Shi, X., Yao, L., 2016. Evaluating energy security of resource-poor economies: a modified principle component analysis approach. Energy Econ. 58, 211–221.

María Marín-Quemada, J., Muñoz-Delgado, B., 2011. Affinity and rivalry: energy relations of the EU. Int. J. Energy Sect. Manage. 5 (1), 11–38.

Martchamadol, J., Kumar, S., 2013. An aggregated energy security performance indicator. Appl. Energy 103, 653–670.

Martchamadol, J., Kumar, S., 2014. The Aggregated Energy Security Performance Indicator (AESPI) at national and provincial level. Appl. Energy 127, 219–238.

Martišauskas, L., Augutis, J., Krikštolaitis, R., 2018. Methodology for energy security assessment considering energy system resilience to disruptions. Energy Strategy Rev 22, 106–118.

Maxim, A., 2014. Sustainability assessment of electricity generation technologies using weighted multi-criteria decision analysis. Energy Policy 65 (C), 284–297.

Molyneaux, L., Wagner, L., Froome, C., Foster, J., 2012. Resilience and electricity systems: a comparative analysis. Energy Policy 47, 188–201.

Mukaka, M.M., 2012. A guide to appropriate use of correlation coefficient in medical research. Malawi Med. J. 24 (3), 69–71.

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.

Nelder, J., 1961. The fitting of a generalization of the logistic curve. Biometrics 17 (1), 89–110.

Nunnally, J.C., Bernstein, I., 1978. Psychometric Theory. MacGraw-Hill. Intentar embellecer nuestras ciudades y también las, New York.

Obadi, S.M., Korcek, M., 2017. EU energy security-multidimensional analysis of 2005–2014 development. Int. J. Energy Econ. Policy 7 (2).

Onamics, 2005. Onamics Energy Security Index—Central and Eastern Europe. Onamics LLC, Washington, DC.

Organization for Security and Co-operation in Europe, Energy Security as a Prerequisite for Political Stability, https://www.osce.org/serbia/16332?download=true.

Ouyang, M., 2014. Review on modeling and simulation of interdependent critical infrastructure systems. Reliab. Eng. Syst. Saf. 121, 43–60.

Pettit, T.J., Fiksel, J., Croxton, K.L., 2010. Ensuring supply chain resilience: development of a conceptual framework. J. Business Logistics 31 (1), 1–21.

Pohekar, S., Ramachandran, M., 2004. Application of multi-criteria decision making to sustainable energy planning—a review. Renew. Sustain. Energy Rev. 8 (4), 365–381.

Prambudia, Y., Nakano, M., 2012. Integrated simulation model for energy security evaluation. Energies 5 (12), 5086–5110.

Radovanović, M., Filipović, S., Pavlović, D., 2017. Energy security measurement – A sustainable approach. Renew. Sustain. Energy Rev. 68, 1020–1032.

Ramanathan, R., 2005. An analysis of energy consumption and carbon dioxide emissions in countries of the Middle East and North Africa. Energy.

Ramezani, H., 2012. A Note on the Normalized Definition of Shannon's Diversity Index in Landscape Pattern Analysis. Environment and. Nat. Resour. Res. 2 (4).

Roege, P.E., Collier, Z.A., Mancillas, J., McDonagh, J.A., Linkov, I., 2014. Metrics for energy resilience. Energy Policy 72, 249–256.

Roupas, C.V., Flamos, A., Psarras, J., 2009. Measurement of EU27 oil vulnerability. Int. J. Energy Sect. Manage. 3 (2), 203–218.

Saisana, M., Philippas, D., 2012. Sustainable Society Index (SSI): taking societies' pulse along social, environmental and economic issues. Environ. Impact Assess. Rev. 32, 94–106.

Saisana, M., Saltelli, A., Tarantola, S., 2005. Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. J. R. Stat. Soc.: Ser. A (Statistics in Society) 168 (2), 307–323.

Saisana, M., Saltelli, A., 2011. Rankings and Ratings: instructions for use. Hague J. Rule Law 3 (2), 247–268.

Saltelli, A., 2007. Composite indicators between analysis and advocacy. Soc. Indic. Res. 81 (1), 65–77.

Sato, M., Kharrazi, A., Nakayama, H., Kraines, S., Yarime, M., 2017. Quantifying the supplier-portfolio diversity of embodied energy: strategic implications for strengthening energy resilience. Energy Policy 105, 41–52.

Scheepers, M., Seebregts, A., de Jong, J., Maters, H., 2007. EU standards for energy security of supply. Gas (6), 52.

Selvakkumaran, S., Limmeechokchai, B., 2013. Energy security and co-benefits of energy efficiency improvement in three Asian countries. Renew. Sustain. Energy Rev. 20, 491–503.

Seo, S., 2006. A Review and Comparison of Methods for Detecting Outliers in Univariate Data Sets. University of Pittsburgh.

Sharifi, A., Yamagata, Y., 2016. Principles and criteria for assessing urban energy resilience: a literature review. Renew. Sustain. Energy Rev. 60, 1654–1677.

Sharifuddin, S., 2014. Methodology for quantitatively assessing the energy security of Malaysia and other Southeast Asian countries. Energy Policy 65, 574–582.

Sheinbaum-Pardo, C., Ruiz-Mendoza, B.J., Rodríguez-Padilla, V., 2012. Mexican energy policy and sustainability indicators. Energy Policy 46, 278–283.

Song, M.-L., Zhang, L.-L., Liu, W., Fisher, R., 2013. Bootstrap-DEA analysis of BRICS' energy efficiency based on small sample data. Appl. Energy 112, 1049–1055.

Sovacool, B.K., 2010. Evaluating energy security in the Asia pacific: towards a more comprehensive approach. Energy Policy 39, 5.

Sovacool, B.K., 2012. The methodological challenges of creating a comprehensive energy security index. Energy Policy 48, 835–840.

Sovacool, B.K., Brown, M.A., 2010. Competing dimensions of energy security: an international perspective. Annu. Rev. Environ. Resour. 35, 77–108.

Sovacool, B.K., Mukherjee, I., 2011. Conceptualizing and measuring energy security: a synthesized approach. Energy 36, 8.

Sovacool, B.K., Mukherjee, I., Durpady, I.M., D'Agostino, A., 2011. Evaluating energy security performance from 1990 to 2010 for eighteen countries. Energy 36, 8.

Spearman, C., 1904. The proof and measurement of association between two things. Am. J. Psychol. 15 (1), 72–101.

Spellerberg, I.F., Fedor, P.J., 2003. A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon–Wiener'Index. Glob. Ecol. Biogeogr. 12 (3), 177–179.

Streiner, D.L., 2003. Starting at the beginning: an introduction to coefficient alpha and internal consistency. J. Pers. Assess. 80 (1), 99–103.

Swiss Re. Sigma Explorer – Catastrophe and Insurance Market Data. 2018 [retrieved on 8 March 2018]; Available from: http://www.sigma-explorer.com.

Tonn, G.L., Czajkowski, J.R., Kunreuther, H.C., 2018. Improving US transportation infrastructure resilience through insurance and incentives. Risk Management and Decision. University of Pennsylvania. Wharton School.

Transmission & Distribution World. India to Spend $21.6 Billion On Smart Grid Infrastructure By 2025. 2019 [retrieved on 5 May 2019]; Available from: https://www.tdworld.com/smart-grid/india-spend-216-billion-smart-grid-infrastructure-

2025.

U.S. Energy Information Administration, Electric Power Monthly with Data for February 2018, 2018, https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t = epmt_6_07_b.

Valdés, J., 2018. Arbitrariness in multidimensional energy security indicators. Ecol. Econ. 145, 263–273.

Velasquez, M., Hester, P.T., 2013. An analysis of multi-criteria decision making methods. Int. J. Oper. Res. 10 (2), 56–66.

Vera, I., Langlois, L., 2007. Energy indicators for sustainable development. Energy 32 (6), 875–882.

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. Int. J. Greenhouse Gas Control 54, 1–14.

Wang, J.-J., Jing, Y.-Y., Zhang, C.-F., Zhao, J.-H., 2009. Review on multi-criteria decision analysis aid in sustainable energy decision-making. Renew. Sustain. Energy Rev. 13 (9), 2263–2278.

Wang, Q., Zhou, K., 2017. A framework for evaluating global national energy security. Appl. Energy 188, 19–31.

Wender, B.A., Morgan, M.G., Holmes, K.J., 2017. Enhancing the resilience of electricity systems. Engineering 3 (5), 580–582.

Willis, H.H., Loa, K., Measuring the Resilience of Energy Distribution Systems, 2015.

Winzer, C., 2012. Conceptualizing energy security. Energy Policy 46, 36–48.

World Bank. Worldwide Governance Indicators – Control of Corruption. 2015 [retrieved on 23 November 2017]; Available from: http://databank.worldbank.org/data/reports.aspx?source = Worldwide-Governance-Indicators.

World Bank. World Governance Indicators. 2015 [retrieved on 29 May 2017]; Available from: http://databank.worldbank.org/data/reports.aspx?source = worldwide-governance-indicators.

World Bank. Worldwide Governance Indicators – Government Effectiveness. 2015 [retrieved on 23 November 2017]; Available from: http://databank.worldbank.org/data/reports.aspx?source = Worldwide-Governance-Indicators.

World Bank. Worldwide Governance Indicators – Political Stability and Absence of Violence/Terrorism. 2015 [retrieved on 23 November 2017]; Available from: http://databank.worldbank.org/data/reports.aspx?source = Worldwide-Governance-Indicators.

World Bank. Ease of doing business index, 2016 [retrieved on 8 March 2018]; Available from: http://data.worldbank.org/indicator/IC.BUS.EASE.XQ.

World Bank. Investment in energy with private participation. 2016 [retrieved on 3 January 2018]; Available from: http://data.worldbank.org/indicator/IE.PPI.ENGY.CD?view = chart.

World Bank, 2017. Distance to Frontier and Ease of Doing Business Ranking. http://www.doingbusiness.org/~/media/WBG/DoingBusiness/Documents/Annual-Reports/English/DB17-Chapters/DB17-DTF-and-DBRankings.pdf.

World Bank. World Development Indicators. 2017 [retrieved on 30 August 2017]; Available from: https://data.worldbank.org/data-catalog/world-development-indicators.

World Economic Forum, Fostering Effective Energy Transition, A Fact-Based Framework to Support Decision-Making, 2018, http://www3.weforum.org/docs/WEF_Fostering_Effective_Energy_Transition_report_2018.pdf.

World Energy Council, 2011 Assessment of country energy and climate policies, 2011, https://www.worldenergy.org/wp-content/uploads/2012/10/PUB_wec_2011_assessment_of_energy_and_climate_policies_2011_WEC.pdf.

World Energy Council, World Energy Trilemma Index, 2018, https://www.worldenergy.org/wp-content/uploads/2018/10/World-Energy-Trilemma-Index-2018.pdf.

Wu, K., Fesharaki, F., Westley, S.B., 2007. Asia's Energy Future: Regional Dynamics and Global Implications. East-West Center, Honolulu, HI.

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.

Wulf, C., Zapp, P., Schreiber, A., Marx, J., Schlör, H., 2017. Lessons learned from a life cycle sustainability assessment of rare earth permanent magnets. J. Ind. Ecol.

Yan, W., Peng, Y., Wang, Z., 2017. Risk assessment for oil & gas resource countries based on PCA-SVM model. 20th International Conference on. Information Fusion (Fusion). IEEE.

Yao, L., Chang, Y., 2014. Energy security in China: a quantitative analysis and policy implications. Energy Policy 67, 595–604.

Zeng, S., Streimikiene, D., Baležentis, T., 2017. Review of and comparative assessment of energy security in Baltic States. Renew. Sustain. Energy Rev. 76, 185–192.

Zhang, X.-P., Cheng, X.-M., Yuan, J.-H., Gao, X.-J., 2011. Total-factor energy efficiency in developing countries. Energy Policy 39 (2), 644–650.

Zhang, H.-Y., Ji, Q., Fan, Y., 2013. An evaluation framework for oil import security based on the supply chain with a case study focused on China. Energy Econ. 38, 87–95.

Zhang, L., Yu, J., Sovacool, B.K., Ren, J., 2017. Measuring energy security performance within China: toward an inter-provincial prospective. Energy 125, 825–836.

Zhou, P., Ang, B.W., 2008. Linear programming models for measuring economy-wide energy efficiency performance. Energy Policy 36 (8), 2911–2916.