Highlights

A comparative sustainability assessment of several grid energy storage technologies

Roque Aguado Molina, Juan José Cartelle Barros, María del Pilar de la Cruz López, Manuel Lara Coira, Alfredo del Caño Gochi

- First application of MIVES–Monte Carlo method in the energy storage sector
- Comprehensive sustainability evaluation of six large-scale energy storage technologies
- Integration of economic, social, environmental, and technical sustainability aspects
- Flow batteries, pumped hydro, and liquid air energy storage have high sustainability indices
- Hydrogen energy storage ranks lowest in overall sustainability performance

A comparative sustainability assessment of several grid energy storage technologies

Roque Aguado Molina[a,*], Juan José Cartelle Barros[b], María del Pilar de la Cruz López[c], Manuel Lara Coira[d], Alfredo del Caño Gochi[e]

[a]*Departamento de Ingeniería Eléctrica, Escuela Politécnica Superior de Linares (EPSL), Universidad de Jaén (UJA), Avda. de la Universidad s/n, 23700 Linares, Spain*
[b]*Escuela Politécnica de Ingeniería de Ferrol (EPEF), Campus Industrial de Ferrol, Centro de Investigación en Tecnologías Navales e Industriales (CITENI), Departamento de Ciencias de la Navegación e Ingeniería Marina, Universidade da Coruña (UDC), Campus de Esteiro, 15403 Ferrol, Spain*
[c]*Escuela Politécnica de Ingeniería de Ferrol (EPEF), Campus Industrial de Ferrol, Departamento de Ingeniería Civil, Universidade da Coruña (UDC), Campus de Esteiro, 15403 Ferrol, Spain*
[d]*Escuela Politécnica de Ingeniería de Ferrol (EPEF), Departamento de Ciencias de la Navegación e Ingeniería Marina, Universidade da Coruña (UDC), Campus de Esteiro, 15403 Ferrol, Spain*
[e]*Escuela Politécnica de Ingeniería de Ferrol (EPEF), Centro de Innovación Tecnológica en Edificación e Ingeniería Civil (CITEEC), Departamento de Ingeniería Civil, Universidade da Coruña (UDC), Campus de Esteiro, 15403 Ferrol, Spain*

Abstract

The global energy transition toward a low-carbon economy is driving increasing penetration of variable energy sources into electricity markets. This unprecedented deployment of intermittent renewables confronts decision-makers in the electricity sector with the challenge of selecting among different energy storage technologies, a choice that must be made on the basis of sustainability criteria. Existing studies present shortcomings, including the absence of the social dimension, the use of weights against sustainable development, or the application of methodologies affected by the rank reversal issue, among others. To address gaps in current knowledge, this study presents a novel probabilistic model for assessing the global sustainability of grid energy storage technologies. The model is based on the MIVES (*Modelo Integrado de Valor para una Evaluación Sostenible*)–Monte Carlo method, which combines requirement trees, value functions, the analytic hierarchy process, and probabilistic simulations. It consists of 19 indicators and makes it possible to obtain a sustainability index (*SI*), as well as partial economic, social, environmental, and technical indices for each technology. Data from an extensive literature review were integrated with expert input and estimations based on linear correlations to address challenges in assessing social and environmental indicators. The model was applied to six technologies: pumped hydroelectric energy storage (PHES), compressed air energy storage (CAES), liquid air energy storage (LAES), vanadium redox flow batteries (VRFB), sodium-sulfur batteries (NaSB), and hydrogen energy storage (HES). A comprehensive sensitivity analysis is also included. To the best of the authors' knowledge, no existing study has utilized the innovative methodology presented in this paper, nor has any related research achieved the scope and depth proposed here. The top-performing technologies identified for the economic, social, environmental, and technical dimensions of sustainability are CAES, VRFB, LAES, and PHES, respectively. In terms of global sustainability, VRFB, LAES and PHES are the best options, while HES consistently ranks last. NaSB and CAES occupy intermediate positions.

Keywords: Electricity storage, Global sustainability, Multi-criteria decision-making, Requirement trees, Value functions, Monte Carlo simulation, MIVES–Monte Carlo method

1. Introduction, literature review and objectives

Ever since the groundbreaking discovery of electricity, there has been an ongoing pursuit to develop effective methods for storing this versatile form of energy [1]. A major advantage of electricity storage is its ability to deliver previously stored energy on demand, providing greater control and flexibility in its utilization. In recent times, the quest for efficient energy storage solutions has gained heightened significance to synchronize power generation and consumption, especially with the rising deployment of low-carbon renewable energy sources and the need to address their inherent intermittency [2]. As a result, nowadays, electricity storage is recognized as one of the four key technical options to provide power system flexibility in low-carbon power systems, alongside flexible generation, network interconnection, and demand-side response [3].

The global power generation landscape is undergoing a significant transformation, driven by the urgent need to mitigate climate change and reduce reliance on fossil fuels. Numerous governmental and intergovernmental organizations, including the International Energy Agency (IEA) and the International Renewable Energy Agency (IRENA), agree in their projections that, in the coming decades, there will be a substantial increase in installed capacity and electricity generation from intermittent renewables, such as wind or solar power [4]. As the share of variable renewable energy sources in the global electricity mix increases, power systems will require enhanced flexibility to accommodate these changes. In this context, energy storage technologies are set to play an increasingly relevant role in ensuring a reliable, sustainable, and resilient energy infrastructure,

*Corresponding author

Email addresses: ramolina@ujaen.es (Roque Aguado Molina), juan.cartelle1@udc.es (Juan José Cartelle Barros), pilar.cruz1@udc.es (María del Pilar de la Cruz López), manuel.lara.coira@udc.es (Manuel Lara Coira), alfredo.cano@udc.es (Alfredo del Caño Gochi)

as well as concurrently delivering cost savings to consumers [5]. A report published by the IRENA in 2017 [6] forecasted that global electricity storage would triple in energy terms, growing from 4.67 TWh in 2017 to 11.89–15.72 TWh in 2030, assuming that the share of variable renewables in the world's power generation mix would double within this timeframe. Similarly, a market report on grid energy storage released by the US Department of Energy in 2020 [7] estimated that the stationary and transportation energy storage combined markets could grow 2.5–4 TWh annually by 2030, representing about three to five times the 800-GWh market size in 2020. More recent data from the IEA indicates that the global installed energy storage capacity is expected to surge from 268 GW in 2023 to 1,016 GW by 2030 in the Stated Policies (STEPS) scenario, with an even greater increase to 1,503 GW by 2030 in the Net Zero Emission (NZE) scenario [8]. At the regional level, other recognized organizations share similar projections for a remarkable expansion of energy storage capacity in the near future. Particularly, according to a 2022 report by the European Association for Storage of Energy (EASE) [9], the energy storage power capacity requirements in the EU are projected to reach approximately 200 GW by 2030, a significant rise from the existing 60 GW. Although ambitious, this target aligns with existing non-binding national objectives set by several EU member states, such as Spain, for example, which aims to increase its energy storage capacity from 8.3 GW in 2021 to around 20 GW by 2030 in its dedicated Energy Storage Strategy [10]. The consensus among all these organizations on the forthcoming growth of energy storage capacity demonstrates its timely and essential role in supporting the global transition to a sustainable energy future.

Energy storage technologies are typically categorized according to the intended purpose for which the energy is stored, being grouped first into electrical and thermal [11]. Electrical energy storage is further subclassified into mechanical (including potential, kinetic, and elastic forms), electrical, chemical, electrochemical, and thermal storage, the latter of which includes cryogenic energy storage [12]. Presently, several storage technologies are available at different stages of development, including pumped hydroelectricity, compressed air, liquid air, flywheels, hydrogen, batteries and flow batteries, supercapacitors, superconducting coils, etc. Each technology has its own characteristics and scale of application, as well as its positive and negative impacts on the economy, society, and the environment. Some of these technologies are currently available, and others are still under development [1]. Fig. 1 provides an overview of the main forms of storable energy and their associated technologies [13].

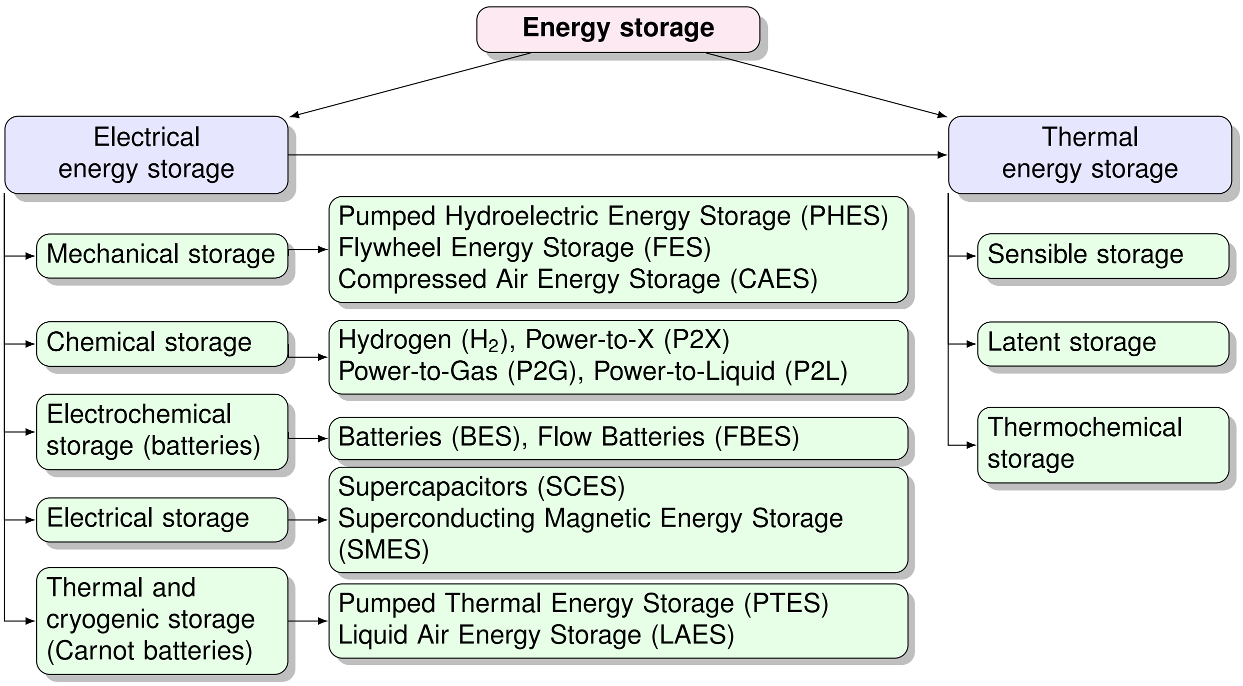

Figure 1: Classification of the main energy storage technologies by the form of stored energy.

Some of the different technologies shown in Fig. 1 can only provide short-term energy storage, while others can last for much longer. Grid-side energy storage refers to large-scale energy storage technologies that are able to operate within an electric power grid, directly interacting with transmission and distribution systems to manage fluctuations in supply and demand. These technologies typically require a minimum rated power of over 10 MW and a capacity-to-power ratio of at least one hour to enable load shifting [14]. Unlike behind-the-meter or distributed energy storage solutions, which are positioned closer to end users, grid-side storage primarily serves bulk power management applications, such as grid stabilization, peak shaving, and frequency regulation. Examples include pumped hydro, compressed air, liquid air, pumped thermal, and hydrogen storage systems, as well as certain types of electrochemical batteries, such as sodium-sulfur (NaS) and flow batteries [11, 14].

Bulk energy storage is currently dominated by pumped hydroelectric energy storage (PHES) in hydroelectric dams [14], a mature technology that currently represents around 96% of the global energy storage capacity [6, 13, 15]. This system stores electrical energy in the form of gravitational potential energy by pumping water between two reservoirs located at different heights [6], connected through pipes, a pump/turbine, and an electric motor/generator. The pump and turbine can either be separate machines or a single device that performs both functions [14], referred to as a reversible pump-turbine [6]. Similarly, the reversible electric generator/motor system can consist of two separate devices (a motor driving the pump and a generator connected to the turbine) or a single electrical machine (an electric motor/generator). The stored energy in a PHES plant is directly related to the volume of water stored in the upper reservoir and the height difference between the reservoirs [6]. PHES technology exhibits very low self-discharge, a reasonable round-trip efficiency, and the capability for substantial volume storage over extended periods [6]. However, a major drawback of this technology lies in its reliance on specific geographical features for installation and susceptibility to political conditions [6]. Despite its large storage volume, PHES has a relatively low energy density, resulting in large land occupation [6]. Moreover, it demands high initial investment costs, a prolonged construction period, and an extended time frame to recover the initial investment [6].

Compressed air energy storage (CAES) is another commercially available technology capable of delivering power outputs above 100 MW, as well as bulk energy storage capacity [16]. CAES systems use electrically driven compressors to store the electrical energy in the form of potential elastic energy of pressurized air in a reservoir [6, 17], which can then be released on demand to generate electricity again by expanding the air through an air turbine. The compressed air is typically stored in underground caverns, predominantly salt caverns, within a pressure range of

4–8 MPa [1, 6]. Similarly to PHES systems, the geographical constraints associated with locating suitable geological formations for compressed air storage are a major drawback of current CAES systems [11, 18]. Despite the significant attention that CAES technology has received in recent years, only two large-scale CAES systems have been successfully connected to the power grid: a 290 MW plant in Huntorf, Germany and a 110 MW facility in McIntosh, Alabama, United States, which were commissioned in 1978 and 1991, respectively [6, 14, 17]. In both facilities, the charge compressor relies solely on electrical energy, while the discharge turbine is supplied with compressed air, supplemented by the use of fossil fuels [14]. Consequently, the traditional CAES configuration operates as a hybrid between a gas turbine power plant and a storage plant. It is thus challenging to compare this technology with other options for grid energy storage [14].

A variant of compressed air energy storage technology belonging to the cryogenic storage category is liquid air energy storage (LAES). This technology is designed for large-scale, long-duration energy storage and uses air or nitrogen (constituting about 78% of ambient air) as the working fluid and storage medium [19, 20]. In LAES systems, instead of directly compressing the air, it is cooled and compressed in multiple stages until it becomes liquefied. Basic air liquefaction systems, such as the Linde cycle, typically result in round-trip efficiencies of less than 40% [14]. By contrast, more advanced systems like the Claude, Kapitza, or Heylandt cycles can reach efficiencies of 50–60% [14, 18]. Even though the Claude cycle is more expensive than the Linde cycle, it is considered the optimal trade-off between performance and cost in air liquefaction [14]. Liquid air is then introduced into cryogenic insulated tanks at low pressure as a means of energy storage, thereby avoiding geographical constraints [14, 18, 21–23]. The system releases stored energy by evaporating the air at high pressure and expanding it in an air turbine, similarly to the process in a Rankine cycle [24]. LAES requires significantly less storage volume compared to CAES, approximately 700 times less [18, 23]. However, LAES technology is not as well known and tested as CAES [25], and is still under industry demonstration level [22]. Moreover, from an economic standpoint, CAES currently provides a shorter payback period and a lower levelized cost of storage [18]. Unlike CAES systems, in which compressed air is the storage medium, LAES systems store electrical energy as thermal energy. LAES is often classified within the Carnot battery category, although this classification may be contentious [26]. Other examples of Carnot batteries include technologies such as pumped thermal electricity storage (PTES), which utilize electric heaters and Rankine or Brayton heat engines. Although these technologies have the potential for grid electricity storage, they have not yet achieved sufficient maturity to be considered competitive alternatives. Indeed, among Carnot batteries, the PTES technology is reported to have a lower degree of maturity than LAES [14]. Therefore, PTES technology has been excluded from this work, not only because of its lower maturity [27, 28], but also due to the scarcity of reliable data. Unlike PTES, a few demonstration-scale LAES systems are already operational and connected to the power grid [22]. Furthermore, several international projects are underway to promote further research and development of LAES technology [23]. The first two pre-commercial facilities were designed, installed, and commissioned in the United Kingdom by the British company Highview Power [20, 22, 23, 27]. A 350 kW/2.5 MWh pilot plant was commissioned in 2011 in Heathrow, UK and a 5 MW/15 MWh demonstration system was commissioned in 2018 in Manchester, UK [28].

Batteries and flow batteries include a wide range of different technologies, some of which can also be used to operate at the electric grid level. In particular, vanadium redox flow batteries (VRFB) and sodium sulfur batteries (NaSB) are usually regarded as large-scale electricity storage technologies [13, 14], since they are often proposed for load-shifting applications, a characteristic that sets them apart from more conventional battery technologies such as Li-ion [14]. VRFBs store energy through internal redox reactions occurring within the cell, driven by active ionic vanadium materials (V^{2+}/V^{3+} and V^{5+}/V^{4+} redox couples), resulting in electron transfer within the circuit [6]. VRFBs are the most widespread and advanced redox flow batteries. They stand out as the only technology that has been deployed in grid-scale applications worldwide for extended periods of time [6, 14]. In contrast to VRFBs, NaSBs operate at high temperatures, typically between 300 and 350 °C, because they rely on sodium (Na) and sulfur (S) molten salts as positive and negative electrodes, respectively [6, 14]. Within each NaS cell, the electrodes are divided by a solid electrolyte composed of β-alumina, which facilitates ionic conduction [6, 14]. NaS cells are usually arranged in series and parallel configurations within a module to achieve the desired power and capacity ratings. NaSB modules ranging from 10 to 50 kW and from 50 to 400 kWh are commercially available [14], which can be combined to reach the megawatt scale [14]. The main advantage of NaSBs lies in their relatively high energy density compared to VRFB technologies, falling within the lower end of the Li-ion energy density range [6]. However, despite their relatively low capital cost and high energy density, NaSBs have not experienced widespread deployment yet, mostly due to concerns related to safety and durability [29].

Hydrogen is an energy carrier that can also be used for bulk storage of electricity. It has been proposed as a means to improve power grid stability by acting as an energy buffer with long-term storage capabilities, thereby balancing supply and demand [30]. This approach, known as power-to-hydrogen-to-power (P2H2P) [30], involves converting surplus electricity into hydrogen, which can then be converted back into electricity, rather than being used for other X applications (P2X), including Power-to-Chemicals (P2C), Power-to-Gas (P2G), Power-to-Liquid (P2L), etc. The process of producing, storing, and re-electrifying hydrogen gas is also termed hydrogen energy storage (HES) [5, 31]. The use of a water electrolysis unit is a common method for producing hydrogen, which can then be compressed or liquefied for storage in high-pressure or cryogenic containers, respectively. Bulk storage of hydrogen can be achieved using pressure vessels or conventional tanks for immediate use, while seasonal hydrogen storage as compressed hydrogen relies on geological formations, such as salt caverns or depleted reservoirs [28, 30]. Nevertheless, despite hydrogen caverns being smaller for a fixed energy content, the cost per kWh of energy discharged from hydrogen caverns can be slightly higher than that from CAES [28]. Liquid hydrogen is not advantageous as a seasonal storage medium, since the lower electricity prices for flexible operation make it challenging to balance the investment costs for liquefaction [30]. For electricity generation from stored hydrogen, a fuel cell, also referred to as a regenerative fuel cell, is generally used [32]. A significant drawback to using hydrogen for electricity storage is the considerable energy loss incurred over a complete charge/discharge cycle [31]. In fact, water electrolysis typically operates at an efficiency of 60%, hydrogen transport and compression for storage may lead to another 10% efficiency loss, while reconversion to electricity through fuel cells has an efficiency of approximately 50%, resulting in an overall round-trip efficiency of around 30% [31]. Hydrogen-based energy storage has been gaining momentum in recent years [33]. However, the P2H2P technology is still in an early stage of entering the electrification sector and has yet to reach the required levels of technological and manufacturing readiness for large-scale deployment [30, 34].

Among the aforementioned technologies for grid energy storage, there is currently no single option that clearly stands out as the most cost-effective, energy efficient, environmentally friendly, and safe. Therefore, it is essential to assess and compare these alternatives in terms of their economic, technical, social, and environmental performance, focusing on their contributions to sustainable development and integral sustainability. The academic community has made some efforts to shed light on this issue. Nevertheless, the scientific literature reveals a scarcity of studies assessing the sustainability of energy storage systems, and those that do exist often present significant limitations. The existing research, along with the gaps in current knowledge, is outlined below and summarized in Table 1.

A noteworthy study in the relevant research is a multi-attribute decision analysis framework consisting of ten criteria distributed in four categories (economic, performance, technological, and environmental) for sustainability assessment of five energy storage technologies, namely pumped hydro, compressed air, lead-acid batteries, lithium-ion batteries, and flywheels [35]. The results were validated using the interval technique for order of preference by similarity to ideal solution (TOPSIS). A sensitivity analysis was also conducted to investigate the effects of the weights of the evaluation criteria on the sustainability ranking of the five energy storage technologies under consideration. The sustainability ranking of the five energy storage technologies demonstrated

Table 1: Summary of existing research on the sustainability assessment of energy storage technologies.

Source	Year	Methods	Technologies considered	Main advantages and limitations
[35]	2018	Multi-attribute decision analysis (MADA) and interval technique for order of preference by similarity to ideal solution (TOPSIS)	PHES, CAES, lead-acid batteries, lithium-ion batteries, and flywheels	• Consideration of 5 technologies and 12 different scenarios. • Absence of social indicators. • Limited environmental impacts.
[36]	2019	Hesitant fuzzy analytic hierarchy process (HFAHP) and hesitant fuzzy technique for order preference by similarity to ideal solution (HFTOPSIS)	PHES, HES, conventional batteries, high-temperature batteries, and flow batteries	• Considerable number of indicators and alternatives. • Weights against the concepts of sustainable development and global sustainability. • Assessment through semantic labels for all indicators. • Use of emissions rather than environmental impact categories.
[37]	2019	Weighted sum method (WSM)	P2G, PHES, and CAES	• Considerable number of indicators. • No social dimension. • Limited approach for environmental impacts.
[38]	2020	Multi-attribute value theory (MAVT)	Conventional and underground PHES	• Considerable number of indicators. • Use of environmental impact categories rather than emissions. • Analysis limited to PHES. • Absence of technical indicators.
[39]	2020	Stepwise weight assessment ratio analysis (SWARA) and additive ratio assessment (ARAS)	PHES, HES, CAES, supercapacitors, thermal energy storage, lead-acid batteries, and lithium-ion batteries	• Large number of technologies and indicators. • All indicators assessed on a scale of 1 to 5. • Allocation of certain indicators to an inappropriate dimension of sustainability. • Environmental dimension limited to CO_2 and greenhouse gas emissions.

a strong dependence on the weights assigned to the ten criteria. Nevertheless, the overarching trend suggested that pumped hydro and compressed air were the most sustainable options for energy storage. Similarly, Acar et al. [36] assessed the sustainability of five different energy storage systems through a hybrid hesitant fuzzy multi-criteria decision-making (MCDM) methodology composed of hesitant fuzzy analytic hierarchy process (HFAHP) and hesitant fuzzy technique for order preference by similarity to ideal solution (HFTOPSIS). The selected energy storage systems were pumped hydro, conventional batteries, high-temperature batteries, flow batteries, and hydrogen. The sustainability assessment methodology considered economic (power cost and energy cost), environmental (pollutant emissions, area requirement, wastewater quality, and solid waste production), social (safety, accessibility, ease of use, and public acceptance), and technical (efficiency, storage capacity, cycling limit, and performance degradation) criteria. The findings suggested that technical performance exerts the highest impact on the contribution to sustainability, whereas social performance exhibits the lowest influence. Moreover, hydrogen demonstrated the highest level of sustainability performance among the selected energy storage alternatives. In the investigation conducted by Vo et al. [37], three technologies to store surplus electricity with high capacity and long discharge time were explored and compared from the perspective of sustainable development. The assessment involved both single-criterion and group multi-criteria analyses. The criteria were categorized into four aspects: technical, economic, social, and environmental criteria. However, social criteria were not assessed as such. This decision was based on the perception that job creation and social benefits were comparable across different storage technologies, and social acceptability was included within the broader scope of general environmental impacts. The technologies considered were P2G, PHES, and CAES. The results revealed that, in terms of sustainable development, P2G was identified as the most suitable option. Guo et al. [38] conducted a comparative study using multi-attribute value theory and scenario analysis to evaluate the overall sustainability performance of conventional and underground pumped hydro energy storage. They considered economic, environmental, and social indicators. The results indicated that conventional pumped hydro energy storage outperformed its underground counterpart in terms of economic and environmental sustainability impact due to economies of scale, while underground pumped hydro energy storage exhibited superior performance in social sustainability impact. Finally, Albawab et al. [39] introduced a sustainability performance index for ranking diverse energy storage technologies using an MCDM model with five main sustainability criteria and seventeen sub-criteria. The selected energy storage technologies were lead-acid batteries, lithium-ion batteries, supercapacitors, hydrogen storage, compressed air energy storage, pumped hydro, and thermal energy storage. Their findings indicated that thermal energy storage and compressed air alternated between first and second place in the rankings. However, it is noteworthy that thermal energy storage operates differently from electrical energy storage systems. Furthermore, certain technologies, such as supercapacitors, may not be suitable for grid energy storage applications.

Despite these valuable insights, most studies on the sustainability of energy storage systems exhibit at least one of the following shortcomings:

- The social dimension of sustainability is often overlooked.
- Indicators that belong to other dimensions of sustainability are incorrectly classified within the social category, leading to potentially misleading interpretations.
- The environmental dimension is usually limited to global warming potential. However, to achieve a more comprehensive sustainability assessment, it is essential to include as many environmental impact categories as possible. Energy storage technologies are no exception to this need.
- Some studies use a set of weights that contradicts the principles of sustainable development and holistic sustainability. Even though minor differences in the importance of sustainability dimensions may be acceptable, significant disparities do not align with the triple bottom line framework.
- Weights are often wrongly defined on the basis of the assumption that all sustainability indicators hold equal importance.
- Certain MCDM methods are susceptible to rank reversal issues when new alternatives are

introduced.

- The comparison is limited to a small number of technologies.
- Quantitative indicators are evaluated using available data based on semantic labels or numerical scales, leading to less accurate results.
- Uncertainty in the assessment is often disregarded, or only considered in the context of defining weights.

Accordingly, the main objective of this study is to address the identified deficiencies by introducing a novel probabilistic model for assessing the sustainability of the main grid energy storage technologies for which sufficient data are available. This model includes a comprehensive set of indicators belonging to the economic, social, technical, and environmental dimensions, which allows for a more in-depth assessment of sustainability than existing models can provide. Qualitative assessments are only conducted for indicators lacking quantitative data. The model is based on the use of the MIVES–Monte Carlo method, a technique that has certain advantages over other commonly used MCDM tools. Namely, MIVES–Monte Carlo is not affected by the rank reversal issue that can arise when new alternatives are added to the comparison. Potential non-linearities in the assessment can be accounted for, incorporating both quantitative and qualitative indicators, integrating the Analytic Hierarchy Process (AHP) for weight assignment, and considering uncertainty in the assessment through a Monte Carlo simulation process. The latter advantage makes it possible to establish distribution functions for input values to the model (including weights and other parameters), resulting in a wide range of outcomes for each alternative rather than a single numerical value. Consequently, a statistical analysis of each technology's sustainability performance can be carried out, exploring all possible real scenarios, since a single energy storage technology can exhibit varying levels of sustainability. To the best of the authors' knowledge, this is the first time that MIVES–Monte Carlo has been applied to rank main grid energy storage technologies according to sustainability, which is another novel aspect of this study beyond those already discussed. In particular, the model is used to compare the sustainability performance of the following six technologies: PHES, CAES, LAES, VFRB, NaSB, and HES. An extensive literature review has been conducted to obtain the data for these technologies, which represents yet another notable contribution of this study. There are no existing sources that compile all the data used to feed the model presented in this paper. This data set may serve as a valuable resource for future research. It is important to note that other grid energy storage technologies with promising potential (for instance, PTES) have been excluded from this work. The reason is twofold: these technologies have not yet reached sufficient maturity, and reliable data on all the considered indicators are currently unavailable. Nevertheless, the model has been designed to be adaptable for future sustainability assessments of other technologies beyond those included in this study. Examples of such additional technologies may include flywheel energy storage (FES) systems, lead-acid batteries, and lithium-ion batteries, which are usually limited to short-duration storage (≤8 h) and therefore not commonly proposed for load-shifting applications [40]. However, cost reductions and technological developments could make them viable options for grid energy storage in the future.

On the other hand, as previously indicated, the massive penetration of intermittent renewable energy sources into the electricity mix is expected to continue in the years to come. Therefore, decision-makers and professionals in the electricity sector will have to decide which storage technologies to opt for and to what extent, carefully considering all their economic, social, environmental, and technological implications, both positive and negative. It is precisely in this task that the model and findings from this study can be of significant help to society, facilitating well-informed decisions that align with integral sustainability. In other words, relying on a single criterion to make decisions (typically focused on minimizing costs) can lead to adverse outcomes, such as the current global environmental and climate crises. Therefore, models like the one presented in this study are essential for guiding decision-makers in avoiding past mistakes and making more balanced, sustainable choices.

The remainder of this article is structured as follows. Section 2 introduces the general methodology, along with the MIVES–Monte Carlo method and detailed descriptions of the indicators and alternatives being examined. The results of the comparative sustainability assessment are presented and discussed in Section 3, in addition to a sensitivity analysis exploring changes in the weights of sustainability indicators, as well as an evaluation of indicator impacts using tornado diagrams. Finally, Section 4 provides concluding remarks and proposes suggestions for future research endeavors.

2. Methodology

The general methodology followed in this study to assess and compare the sustainability performance of various grid-scale energy storage technologies is outlined in Fig. 2. The first step involved conducting an extensive literature review with a fourfold purpose: i) identifying the main gaps in current knowledge, ii) drawing up a comprehensive list of potential sustainability indicators, iii) collecting data to feed the model, and iv) selecting the technologies to be assessed. After that, the model was defined according to the MIVES–Monte Carlo method. A requirement tree was developed with various levels of disaggregation, establishing specific input values for each indicator across all alternatives. Value functions and weights were determined, and distribution functions were used where applicable. The Delphi technique was used to assist in the process of defining the model's parameters. It is noteworthy that the model's input values for the different alternatives represent a partial result of this study, as there are no sources in the scientific literature providing data for all the sustainability indicators addressed herein. This is the main reason why the model inputs are included in Section 3. The same applies to the parameters (weights and value functions) used to configure the model. In the third phase, the model was executed through a Monte Carlo simulation process until convergence was reached, resulting in a comprehensive set of total and partial sustainability indices for each technology. From these values, a statistical analysis was performed, and the corresponding cumulative probability curves and frequency histograms were generated. Finally, a sensitivity analysis was performed to evaluate the impact of different model parameters. Additional details about the procedure followed are provided in the following subsections.

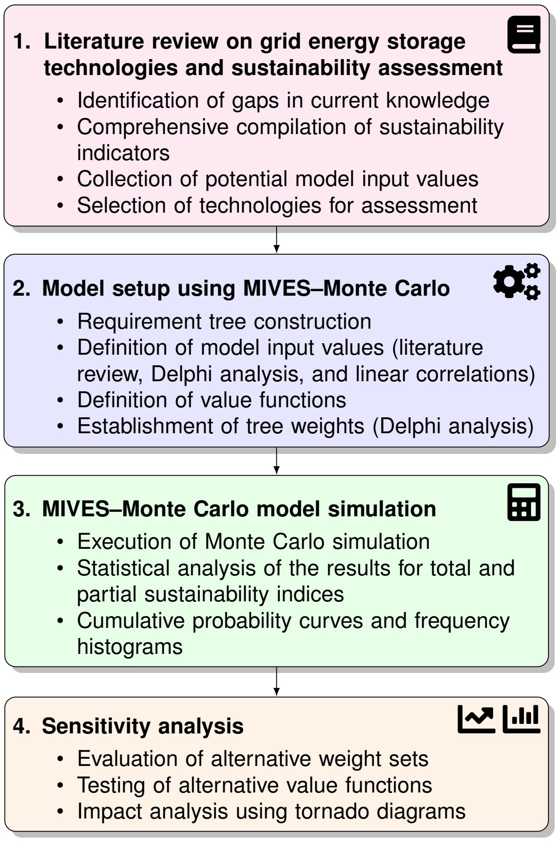

Figure 2: Methodological flowchart for the sustainability assessment of several grid energy storage technologies using the MIVES–Monte Carlo approach.

2.1. MIVES–Monte Carlo method

MIVES, an acronym in Spanish for *Modelo Integrado de Valor para una Evaluación Sostenible* (translated as Integrated Value Model for Sustainability Assessment), is a multi-criteria decision-support method, initially conceived for sustainability assessment. MIVES models are structured as requirement trees, with sustainability indicators branching on the right side and a sustainability index located on the left side as the key variable to be assessed. Partial indices for the environmental, social, economic, and technical dimensions of sustainability are typically evaluated as well. The requirement tree can include multiple intermediate levels to effectively represent the complexity of the problem and simplify the assignment of weights. Nevertheless, requirement trees are generally organized into three levels: requirements, criteria, and indicators.

MIVES makes it possible to use qualitative and quantitative indicators. To perform calculations with these indicators, MIVES relies on value functions to transform them into a dimensionless unit, usually called satisfaction index or satisfaction level, which takes values between 0 and 1, the worst and best possible results, respectively. MIVES also incorporates a set of methods to establish the relative weights of the different indicators, criteria, and requirements. The assessment is made by means of a weighted summation for the whole tree. Finally, combined with non-deterministic techniques, MIVES allows accounting for the uncertainty that may affect the different model parameters specific to each problem. In this study, the Monte Carlo simulation has been combined with MIVES, resulting in the MIVES–Monte Carlo technique. Further information on the MIVES–Monte Carlo method can be found in the scientific literature [41]. Fig. 3 presents a flowchart outlining the main steps of the MIVES and MIVES–Monte Carlo methods.

In the field of sustainability assessment, MIVES has been applied to structural systems [42–47], urban infrastructures [48], buildings and building subsystems [49–52], urban planning and development [53], and power plants, in all cases obtaining valuable results [54–56]. In particular, Aguado et al. [42] introduced a MIVES model for assessing the sustainability of concrete structures. This model has been part of the Spanish Structural Code EHE-08 and aims to optimize the consumption of raw materials, extend their useful life, and promote the use of materials obtained through energy-efficient and low-polluting processes when designing a concrete structure. Pons et al. [43] developed a MIVES model to estimate the sustainability index of concrete pile foundations. It was applied to a case study in which the authors compared reinforced concrete, steel fiber-reinforced concrete, and polymeric fiber-reinforced concrete piles. According to the results, the inclusion of alternative materials serves to enhance the contribution to sustainable development. A MIVES model for assessing the sustainability of structural systems was also presented by de la Fuente et al. [44]. In this case, the authors focused on concrete slabs. They compared

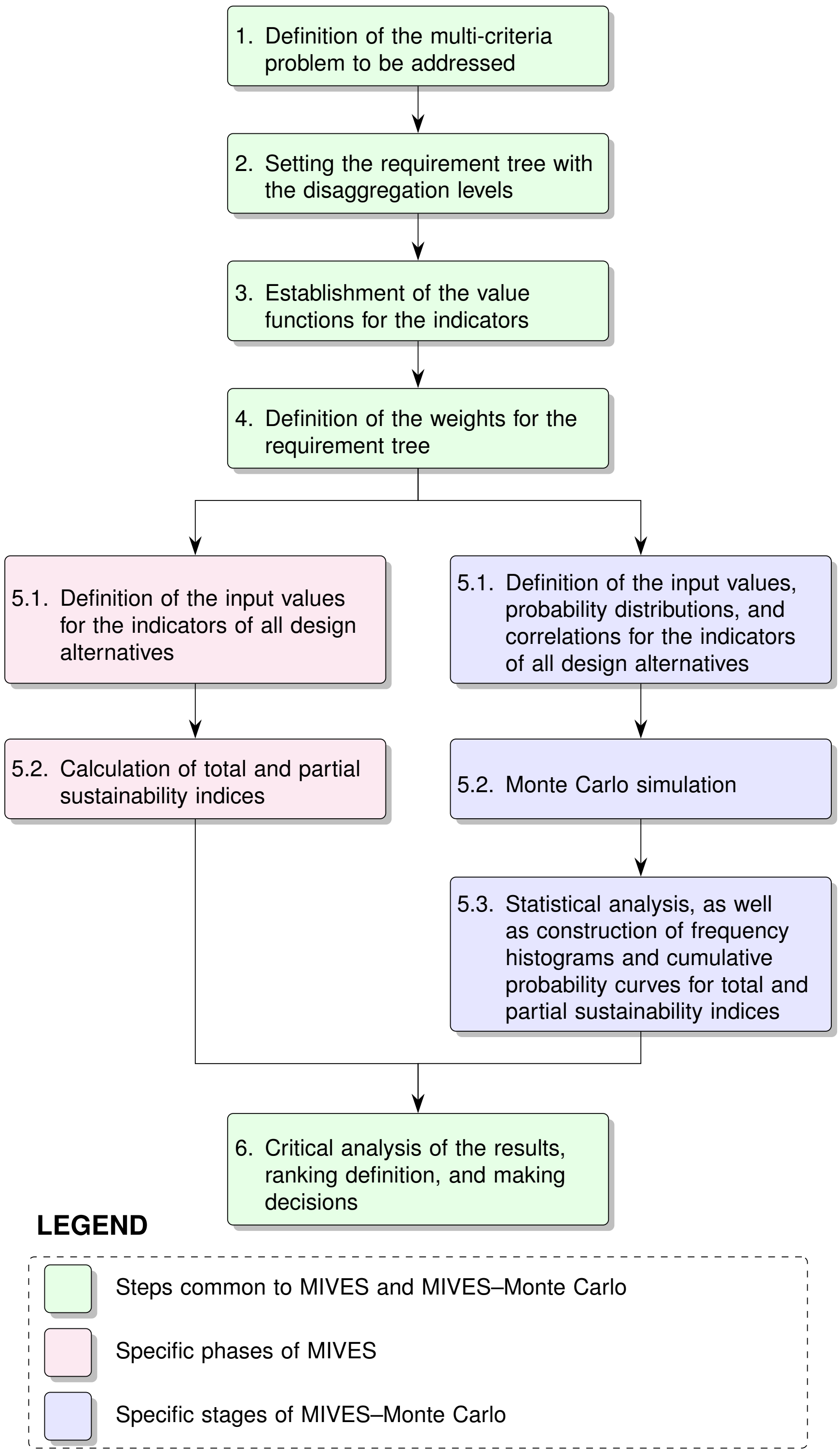

Figure 3: Flowchart of the MIVES and MIVES–Monte Carlo methods.

reinforced concrete and steel fiber-reinforced concrete alternatives for an office building located in Guipúzcoa, Spain. Josa et al. [45] addressed the case of girders and trusses. The authors compared different materials and structure types for a sport center in Catalonia, Spain. They concluded that timber and concrete trusses, despite being used much less in practice, are more sustainable than steel trusses. The environmental dimension of timber structures was studied by Cuadrado et al. [46]. The authors considered three different case studies in which they have changed the type and origin of the timber products for the same structure. The authors demonstrated that the origin of the timber, and whether or not it is certified, can significantly affect the environmental performance of a structure. The same model was later applied to a more complex case study (a real residential building) [47], leading to similar conclusions. MIVES was also valuable for urban infrastructures. In particular, Casanovas-Rubio et al. [48] proposed a model for evaluating the sustainability of narrow trenches for service pipes. They analyzed the same application with four different trenches: standard solution, standard solution with recycling, standard solution with a controlled low-strength material replacing compacted soil, and a new alternative called eco-trench. Significant differences in terms of sustainability were found by the authors among these four options. Buildings have also been a field of application for MIVES. For example, San-José Lombera and Cuadrado [49] addressed the environmental dimension of an industrial building for graphic design and printing activities through a MIVES model made up of indicators belonging to the following thematic areas: location, consumption of energy and water, materials, waste, and construction and deconstruction processes. They compared conventional practices with the most environmentally friendly processes, obtaining a difference close to 60% in the overall index. Subsequently, the sustainability of industrial buildings was extended to include social and economic indicators [50]. MIVES was also applied to school centers. This is the case of the study carried out by Pons and Aguado [51]. They compared prefabricated and non-prefabricated technologies through a MIVES model consisting of 17 indicators. Prefabricated concrete and prefabricated steel turned out to be the most sustainable options. MIVES can also be applied to building subsystems.

For example, Casanovas-Rubio and Armengou [52] created a model for water heating systems. Its validity was tested by comparing 12 options for a dressing room with capacity for 100 people in Barcelona, Spain. MIVES was also used to facilitate decision-making in urban planning. In fact, Herranz-Pascual et al. [53] proposed a model with 50 indicators for evaluating the sustainability of municipal action plans. As for the energy sector, Cartelle Barros et al. [54] developed a MIVES model made up of 27 indicators to compare the contribution to sustainability of renewable and non-renewable power plants. From the results, it is clear that renewables are usually more sustainable, with the exception of biomass power plants. The economic dimension of the main power generation technologies was subsequently analyzed in depth by the same authors [55]. In this case, two probabilistic models were employed, one of them based on the MIVES–Monte Carlo method. This study showed that some renewables could compete with non-renewables in terms of cost. MIVES–Monte Carlo was also successfully used to compare the environmental behavior of power plants. Specifically, Cartelle Barros et al. [56] presented a probabilistic model that incorporates 15 environmental impact indicators with cradle-to-grave life cycles and global geographical scopes. It was used to compare 11 technologies, with wind and hydro standing out as the best-performing alternatives.

Despite having already been used in the broader context of the energy sector, this work represents the first application of MIVES to energy storage. The complex nature of energy storage systems, which involve social, economic, environmental, and technical impacts, makes MIVES especially relevant for assessing their sustainability. MIVES enables the adimensionalization and integration of a wide range of indicators measured in different units and with varying degrees of importance. The value functions in MIVES do not necessarily maintain proportionality during adimensionalization (e.g., doubling the input value does not always correspond to a doubling of satisfaction for increasing indicators, nor does halving it always correspond to double satisfaction for decreasing indicators). This flexibility makes MIVES especially effective for integrating diverse indicators in the sustainability assessment of energy storage systems. One key advantage of

MIVES is that it is not affected by the rank reversal issue when additional alternatives are incorporated into the assessment. As the development of new energy storage technologies continues and more data become available, it is highly desirable that the sustainability ranking of previously assessed technologies is not affected each time a new technology is included. Another important aspect is that MIVES can be easily combined with Monte Carlo simulations to account for the inherent variability and uncertainty of energy storage technologies, particularly those with a lower TRL.

2.2. Requirement tree and sustainability indicators

As illustrated schematically in Fig. 4, the MCDM method proposed in this paper involves using a requirement tree with three levels: requirements, criteria, and indicators. At the first level, requirements are classified into four categories: economic, environmental, social, and technical. This classification has been adopted in similar studies that examine the sustainability performance of energy storage technologies [36]. However, other studies often omit certain aspects, particularly social criteria, due to the inherent complexity of their quantitative assessment [35, 37]. The intermediate levels include the criteria, which aid in breaking down the requirement tree and facilitate the conceptualization and calculation processes of the model. At the third level are the indicators, which define the specific properties to be assessed. A structured overview of all the indicators considered in this study, organized by category, is presented in the following subsections.

2.2.1. Economic indicators

The economic indicators consist of five widely used metrics in energy storage systems. A concise description of each one is provided below.

- Levelized cost of storage (LCOS). Ratio between total expenditure and amount of electrical energy retrieved from the storage system during its lifetime [26, 57], expressed in \$/kWh. This parameter is similar to the well-known levelized cost of electricity (LCOE) used in power generation systems, but modified for the case of electricity storage. The LCOS can

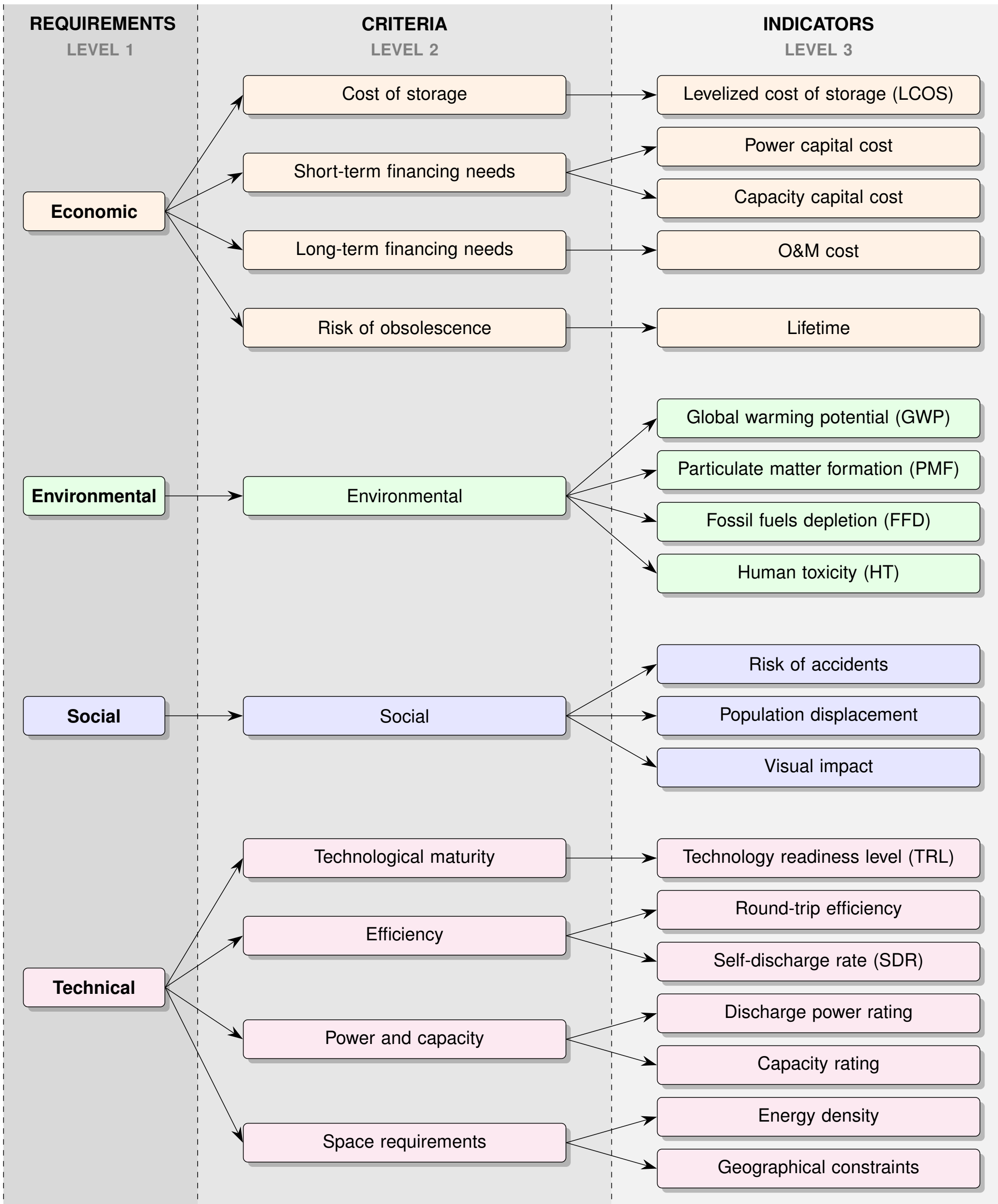

Figure 4: Multi-criteria decision-making (MCDM) framework for evaluating energy storage systems based on a requirement tree integrating economic, environmental, social, and technical aspects.

be interpreted as the average $/kWh price at which energy output from the storage system would need to be sold to break even on total costs [28].

$$\text{LCOS} = \frac{\text{CAPEX} + \sum_{t=1}^{t=N} \dfrac{A_t}{(1+i)^t}}{\sum_{t=1}^{t=N} \dfrac{W_{discharge}}{(1+i)^t}} \tag{1}$$

where A_t denotes the annual cost of the storage system. A_t is composed of the operation cost (OPEX_t), the reinvestments in storage system components ($\text{CAPEX}_{re,t}$) at time t, as well as the cost of electricity supply, which is determined by multiplying the unit price of electricity (c_e) with the annual electricity input (W_{charge}) [57]. At the end of the storage lifetime ($t = N$), a recovery value (R_t) is included. As the cost of electricity (c_e) typically fluctuates over time, an average value may be used.

$$A_t = \text{OPEX}_t + \text{CAPEX}_{re,t} + c_e\, W_{charge} - R_t \tag{2}$$

- Power capital cost (Γ_P). Ratio between capital expenditure (CAPEX) and electrical power output [26], expressed in \$/kW. This indicator assesses the challenges in securing the short-term financing required for the energy storage system relative to its power rating.

$$\Gamma_P = \frac{\text{CAPEX}}{\dot{W}_{discharge}} \tag{3}$$

- Capacity capital cost (Γ_C). Ratio between capital expenditure (CAPEX) and electricity storage capacity [26], expressed in \$/kWh. This indicator evaluates the short-term financing requirements of an energy storage system relative to its storage capacity.

$$\Gamma_C = \frac{\text{CAPEX}}{W_{discharge}} \tag{4}$$

- O&M cost. Operating and maintenance costs (OPEX) associated with the energy storage system, expressed in \$/(kWh·yr). This indicator assesses the challenges of securing long-term financing by evaluating the ongoing costs required for operation and maintenance, which can impact the overall financial viability of the investment.

- Lifetime (*N*). The operational lifespan of the energy storage system, expressed in years (yr). Although a long service life is typically seen as beneficial, it can also lead to premature obsolescence, especially in fields with rapid innovation and development, such as the energy storage sector. In this context, a longer lifespan may hinder the timely adoption of newer, more advanced, and sustainable technologies, with adverse effects on the profitability of the investment, as replacing outdated systems becomes more challenging. For this reason, unlike conventional evaluations, this indicator is negatively assessed. In other words, a longer service life corresponds to a lower satisfaction.

2.2.2. *Environmental indicators*

The environmental indicators consist of four mid-point environmental impact categories. The most widely available environmental indicator in grid energy storage systems is the global warming potential. However, three other indicators, for which data are available in the scientific literature, have also been considered. These environmental indicators have been sourced from life cycle assessment studies with a "cradle-to-grave" scope, where the functional unit is one kWh of electricity delivered back to the grid. Each indicator is described briefly below.

- Global warming potential (GWP). Measure of the consequences of the rise in the planet's temperature, caused by the greenhouse gas emissions associated with the energy storage system, expressed in g CO_2-eq./kWh.
- Particulate matter formation (PMF). Assessment of the production of particulate matter by the energy storage system, expressed in g PM_{10}-eq./kWh.
- Fossil fuels depletion (FFD). Evaluation of the impact of the energy storage system on fossil fuel resources, expressed in g oil-eq./kWh.
- Human toxicity (HT). Assessment of the potential harm to human health by the energy

storage system, expressed in g 1,4-DB-eq./kWh.

2.2.3. *Social indicators*

Social indicators include the risk of accidents, population displacement, and visual impact. These qualitative indicators have also been considered in earlier research works on the sustainability assessment of energy systems [54]. However, it is important to note that previous studies are specifically focused on evaluating the sustainability of power plants, rather than large-scale energy storage technologies. Although these social indicators remain conceptually relevant for energy storage, the input data and assessment methods for these indicators are not directly applicable in this context. Unfortunately, quantitative values for social indicators related to energy storage technologies were unavailable for this study. As a result, these indicators were evaluated qualitatively. A short description of each one follows.

- Risk of accidents. Qualitative evaluation of the likelihood and consequences of accidents related to the energy storage system.
- Population displacement. Qualitative assessment of the impact on human populations due to the installation or operation of the energy storage system. For instance, PHES facilities may require the relocation of residents from communities that become submerged due to the construction and operation of these energy storage systems.
- Visual impact. Qualitative evaluation of the aesthetic effects of the energy storage system on its surroundings.

In addition to the aforementioned indicators, it would have been desirable to consider other social parameters, such as social acceptance, employment generation, and employee satisfaction [58–60], among others. However, for the energy storage technologies included in this study, not only are quantitative data on these additional social indicators missing, but qualitative assessments are also remarkably challenging, especially for less mature technologies. Consequently, these parameters were not evaluated, although they should be incorporated into future versions of the model as soon as reliable data become available.

2.2.4. Technical indicators

Numerous technical indicators are documented in the scientific literature. However, only seven have been selected for inclusion, as they have been identified as the most relevant for sustainability assessment. A brief overview of each one is outlined below.

- Technology readiness level (TRL). An estimate for the maturity of the energy storage technology, reflecting its potential for short, medium, or long-term utilization. TRLs are based on a scale from 1 to 9, with 9 being the most mature technology [61].
- Round-trip efficiency (η_{rt}). Ratio between delivered and absorbed electricity, over a complete charge/discharge cycle [26]. Typically, this efficiency is expressed on an AC-to-AC basis, incorporating any losses from AC-to-DC or DC-to-AC power conversions, if applicable [5]. The round-trip efficiency of a storage system can be broken down into the efficiencies of each stage of the cycle, as shown in Eq. (5).

$$\eta_{rt} = \eta_{charge}\,\eta_{storage}\,\eta_{discharge} = \frac{E_{storage}}{W_{charge}}\,\frac{E'_{storage}}{E_{storage}}\,\frac{W_{discharge}}{E'_{storage}} = \frac{W_{discharge}}{W_{charge}} \tag{5}$$

where η_{charge}, $\eta_{storage}$, and $\eta_{discharge}$ correspondingly denote the efficiencies of the charge, storage, and discharge stages that make up the round-trip cycle [62]. W_{charge} and $W_{discharge}$ represent the energy absorbed during charging and the energy delivered during discharging, respectively. $E_{storage}$ denotes the energy stored immediately after charging. The energy remaining after a storage period from t_0 to t_f, accounting for any losses during this interval, is designated as $E'_{storage}$ and can be mathematically expressed as follows:

$$E'_{storage} = E_{storage} - \int_{t_0}^{t_f} \dot{E}_{sd}\,\mathrm{d}t \tag{6}$$

- Self-discharge ($\dot{E}_{sd}$). The rate at which the storage system loses stored energy over time.

$$\dot{E}_{sd} = -\frac{\mathrm{d}E_{storage}}{\mathrm{d}t} \tag{7}$$

It should be noted that $\frac{\mathrm{d}E_{storage}}{\mathrm{d}t}$ is usually a negative value, since $E_{storage}$ decreases over time due to self-discharge. Accordingly, Eq. (7) incorporates a negative sign to ensure that the self-discharge rate is expressed as a positive value. Self-discharge is often expressed as a fraction or percentage of the energy initially accumulated after charging per unit time, typically days in the context of grid energy storage systems. Thus, the self-discharge rate (SDR) is defined as the ratio of $\dot{E}_{sd}$ to $E_{storage}$, and its unit corresponds to those of the reciprocal of time.

$$\mathrm{SDR} = \frac{\dot{E}_{sd}}{E_{storage}} \tag{8}$$

- Discharge power rating ($\dot{W}_{\max}$). The maximum power output of the energy storage system, expressed in MW.
- Capacity rating ($W_{\max}$). The maximum capacity of the energy storage system, expressed in MW.
- Energy density (ρ_E). Ratio between the electricity released over a complete discharge and the total storage volume [26], expressed in kWh/m^3.

$$\rho_E = \frac{W_{discharge}}{V_{storage}} \tag{9}$$

- Geographical constraints. Limitations imposed by the geographical location on the installation or operation of the energy storage system. For example, PHES facilities need substantial space and water resources, while CAES systems require specific geological features, such as salt caverns.

2.2.5. *Compilation of quantitative indicators*

Table 2 compiles the typical values or ranges of all the quantitative indicators, including references, along with their corresponding units of measurement. The only categorical indicators included in Table 2 are TRL, which takes discrete values ranging from 1 to 9, and geographical constraints, which is a binary parameter that only takes two values: yes and no. Both categorical indicators are well-documented in the scientific literature for each of the grid storage technologies included in the comparative sustainability assessment. By contrast, the social requirement and its corresponding indicators are not featured in Table 2 due to the absence of reliable quantitative values and references. Quantifying the social dimension of sustainability is challenging [54, 58]. Therefore, social indicators have been assessed qualitatively using categorical units, incorporating expert input through the Delphi method.

Table 2: Grid energy storage technologies and quantitative indicators considered for the comparative sustainability assessment. Not available data is as denoted as N/A.

Requirement	Indicator	Unit	Grid energy storage technology					
			PHES	CAES	LAES	NaSB	VRFB	HES
Economic	Levelized cost of storage (LCOS)	$/kWh	0.02–0.31 [63] 0.07–0.14 [64] 0.06–0.11 [57]	0.02–0.16 [63] 0.06–0.13 [64] 0.23–0.28 [65]	0.14–0.46 [66]	0.29–0.34 [67]	0.32–0.56 [67] 0.20–0.30 [28]	0.30–0.50 [57] 0.45–0.59 [68] 0.44–0.54 [69]
	Power capital cost (Γ_P)	$/kW	600–2000 [11, 12, 31, 70] 500–1500 [71, 72]	400–800 [11, 12, 31, 70] 400–1000 [14, 73]	900–2000 [11, 14, 70] 100–2000 [31] 700–3000 [26] 300–1000 [73, 74]	1000–3000 [1, 12, 32, 70] 150–3000 [12] 350–3000 [32] 150–900 [14]	600–1500 [12, 31, 32, 70, 72] 175–1500 [12] 300–1500 [14]	1900–6300 [70] 1900–10,000 [11] 1000–3000 [28] 2400–4700 [74]
	Capacity capital cost (Γ_C)	$/kWh	5–100 [1, 11, 12, 14, 31, 32, 70, 73, 74] 10–20 [71, 72]	2–50 [11, 12, 31, 70, 72] 2–120 [72], 1–200 [14] 4–250 [20], 2–250 [73]	260–530 [11, 14, 32, 70] 240–640 [20], 400–800 [26] 100–300 [27]	300–500 [1, 12, 32, 70] 250–500 [12] 100–600 [14]	150–1000 [12, 14, 31, 32, 70, 72]	1–10 [70] 45–300 [28]
	O&M cost	$/(kW·yr)	2–10 [67], ~ 3 [32] 15–52 [28]	2–5 [67], 10–20 [28] 19–25 [32]	13.5–30 [28]	2–54 [67] ~ 80 [32]	4–51 [67], ~ 6.7 [28] ~ 70 [32]	17–27 [28], 17–48 [67] 8–24 [28]
	Lifetime (N)	yr	30–60 [11, 70, 73, 74] 40–60 [6, 12, 20, 31, 32] 50 [28], 75 [23]	20–40 [12, 20, 70, 73–75] 20–60 [11, 31]	20–40 [11, 14, 32, 70, 73, 74] 30–40 [20]	10–15 [1, 12, 32, 70]	5–15 [11, 70] 5–10 [1, 12, 31, 32] 20–25 [28]	20–30 [70], ~ 30 [28] 5–30 [11] 5–20 [74]
Environmental	Global warming potential (GWP)	g CO_2-eq./kWh	23.5–650 [76] 58–530 [77] 145–179 [78]	27.1–740 [76] 292 [79], 380 [80] 161–272 [78]	30.2–203 [81]	37.9–640 [76] 201–937 [78]	52.8–433 [81] 21–279 [82]	50.6–1620 [76] 386–700 [78]
	Particulate matter formation (PMF)	g PM_{10}-eq./kWh	0.05–0.85 [76]	0.08–1.00 [76]	N/A	0.08–0.85 [76]	N/A	0.16–2.18 [76]
	Fossil fuels depletion (FFD)	g oil-eq./kWh	5.61–189 [76]	7.48–217 [76]	N/A	9.97–187 [76]	N/A	17.4–475 [76]
	Human toxicity (HT)	g 1,4-DB-eq./kWh	15.7–418 [76] 128.3 [38]	24.9–475 [76]	N/A	15.3–417 [76]	44–150 [82]	35.3–1030 [76]
Technical performance	Technology readiness level (TRL)	1–9	9 [14, 27, 70]	7–8 [14, 70], 9 [27]	5–6 [14, 66, 70], 7 [27]	9 [14, 70]	9 [14, 27, 70]	4–5 [70]
	Round-trip efficiency (η_{rt})	%	65–85 [11, 14, 31, 70] 75–80 [67], 70–80 [23, 71] 70–85 [6, 32], 65–87 [11] 75–85 [32, 83], 80 [28]	40–60 [20, 70], 40–70 [74] 42–70 [73], 54–70 [75] 45–70 [83], 40–75 [67, 84] 40–89 [11], 50–89 [31] 40–95 [11]	40–60 [26], 45–60 [20] 30–70 [83], 45–70 [27, 73] 50–70 [23, 31] 50–60 [14, 81] 55–80 [32, 75], 40–85 [11]	75–90 [12, 32, 70], 70–90 [12, 14] 75–85 [32] 70–85 [67, 84]	57–85 [11], 60–85 [12, 14, 70] 60–75 [67, 84] 65–75 [32] 75–85 [1, 32]	30–50 [70] 20–50 [11] 20–66 [74] 35–40 [84] ~ 31 [28]
	Self-discharge rate (SDR)	%/day	0.005–0.02 [6, 11, 70]	0.003–0.03 [11, 70]	~0 [70]	~20 [12] 0.05–20 [70]	~0.20 [70]	~0 [11]
	Discharge power rating	MW	100–5000 [1, 11, 12, 20, 31, 32, 67, 70] 1000–3000 [67]	5–300 [11, 12, 31, 67, 70] 1–320 [20, 73] 10–320 [27]	10–100 [11, 18, 70] 10–200 [32] 20–100 [26] 1–300 [20, 27, 73]	0.05–8 [12, 67] 0.05–50 [14, 70] 0.05–34 [67]	<3 [12, 31] ~12 [67] 0.05–7 [70] 1–25 [14]	0.1–1000 [11, 70] 0.1–50 [67]
	Capacity rating	MWh	> 1000 [11, 14, 70] 500–8000 [1, 32]	> 1000 [11, 70] 580–2860 [32, 74, 75] 640–2640 [27]	>1000 [14] 10–1000 [11, 70] 40–400 [26]	6–600 [70] 1–250 [14]	0.05–10 [70] 1–100 [14]	100–1000 [11, 70]
	Energy density (ρ_E)	kWh/m^3	0.5–1.5 [1, 11, 12, 14, 20, 31, 32, 70, 73] 0.2–2 [23] 0.5–2 [21, 74]	2–6 [21, 74, 75] 3–6 [12, 31, 32] 3–12 [11, 70], 3–20 [20] 0.5–20 [14, 23, 27, 73]	~50 [11, 14, 70], 55 [26] 50–80 [27], 20–110 [26] 120–200 [20], 32–230 [85] 60–200 [23], 50–200 [73] 150–250 [21]	150–250 [12, 14, 32, 70] 150–240 [31]	16–33 [70] 10–30 [31], 16–60 [11, 14] 10–60 [27]	500–3000 [11, 32, 70, 74]
	Geographical constraints	Yes/No	Yes [11, 14, 20, 26, 73, 83]	Yes [11, 14, 20, 26, 27, 73, 83]	No [11, 14, 20, 21, 26, 27, 73]	No [11]	No [11, 26]	Yes [28], No [11, 83]

2.3. Sustainability index and value functions

The sustainability index, denoted as *SI*, is the global performance metric used for the comparative sustainability assessment. It is calculated as the weighted sum of value function outputs for each indicator, as given by Eq. (10) [41, 54].

$$SI = \sum_{i=1}^{i=k} \alpha_i \beta_i \gamma_i V(P_i) \tag{10}$$

where k refers to the number of indicators in the requirement tree; α_i, β_i, and γ_i represent the weights assigned to the requirements, criteria, and indicators, respectively; and $V(P_i)$ denotes the output of the value functions, which normalizes the units of measurement across indicators to a scale from zero to one, where the extreme values correspond to the minimum and maximum levels of satisfaction. The value functions also facilitate the incorporation of potential non-linearities in the sustainability assessment through diverse function shapes, allowing for different levels of exigency in meeting the sustainability requirements for each specific indicator [54]. Furthermore, Eq. (10) can be applied to a subset of criteria and indicators belonging to a single requirement, generating a partial sustainability index (economic, environmental, social, or technical).

The value functions for each continuous quantitative indicator are defined according to Eq. (11).

$$V(P_i) = \frac{1 - \exp\left\{\left[-m_i \left(\frac{|P_i - P_{i,\min}|}{n_i}\right)^{a_i}\right]\right\}}{1 - \exp\left\{\left[-m_i \left(\frac{|P_{i,\max} - P_{i,\min}|}{n_i}\right)^{a_i}\right]\right\}} \tag{11}$$

where P_i is the input value of the indicator i for the alternative under assessment. Furthermore, $P_{i,\min}$ and $P_{i,\max}$ denote, respectively, the values of P_i corresponding to the minimum and maximum levels of satisfaction, i.e., 0 and 1. The terms a_i, n_i, and m_i are shape parameters used to produce concave, convex, S-shaped, or linear value functions.

Unlike continuous quantitative indicators, both qualitative and discrete quantitative indicators require stepped functions, where each step corresponds to a particular category or semantic label. Each category is in turn associated with a satisfaction level ranging from 0 to 1 through qualitative value functions. Therefore, the use of semantic labels does not necessarily limit the accuracy or sensitivity in assessing and comparing the performance of different technologies for a given indicator, since each of them is associated with a specific numerical value.

2.4. Assignment of weights

The next step in the MIVES–Monte Carlo methodology involves establishing the weights of the requirements (α_i), criteria (β_i) and indicators (γ_i). Some authors assign equal importance to all aspects of the evaluation [86]. However, this is not advisable, as it may lead to inaccurate assessments [87]. All requirements contribute to both sustainable development and global sustainability, and thus, excessively large differences among them should be avoided [54].

The Delphi method can be used to determine the weights, with the collaboration of experts in defining indicators, the tree structure, as well as social indicators. The Delphi method operates on the premise that predictions (or decisions) derived from a structured group of individuals are more reliable compared to those from unstructured groups. In this approach, experts participate in successive rounds of questionnaire-based surveys. Following each round, a facilitator presents an anonymized synthesis of the experts' predictions from the preceding round, along with the rationales behind their judgments. Thereafter, experts are prompted to review their initial responses in light of the feedback received from other panel members. This iterative process is expected to narrow the range of responses and facilitate the group's convergence toward a certain answer. Further information about the Delphi method can be found in the scientific literature [88–93].

The assignment of numerical values to the weights of requirements, criteria, and indicators can be rather straightforward in some cases. Indeed, directly allocating weights to branches containing up to four indicators typically presents no significant challenges [54]. However, in certain situations where the requirement tree is overly intricate or when discrepancies among experts occur, it is recommended to adopt a more rigorous methodological approach, such as AHP [94]. Moreover, conducting a subsequent analysis and comparison process is encouraged to mitigate the inherent subjectivity. As a result, adjustments to the corresponding weights can be made, if necessary.

3. Results and discussion

This section delves into the results of the sustainability assessment for the six different grid energy storage technologies. It begins with an overview of model inputs, including the use of triangular and uniform distributions, and addressing gaps in data through estimation and expert-based semantic labeling. The requirement tree and weights used for the model are described thereafter, followed by a presentation of the sustainability assessment results from Monte Carlo simulations, including the sustainability indices and cumulative probability curves for the different technologies. Subsequently, the results from the comparative sustainability assessment are discussed and contrasted with existing research to underscore the unique features of the model. Finally, a sensitivity analysis is conducted to evaluate the impact of variations in the model inputs on the overall results.

3.1. Model inputs

Table 3 presents the input values for the model, assessed individually for each alternative. These values (or ranges of values) are considered as representative of the data previously reported in Table 2 and are best understood in terms of orders of magnitude rather than precise numerical figures. Modal values have also been incorporated into the model when available, enabling the use of both triangular and uniform distributions. Although the results obtained using these two distributions do not show substantial differences, triangular distributions are more representative because modal values are considered in the calculation process [56].

Unfortunately, complete information on all energy storage systems is unavailable for certain environmental impact indicators. Particularly, reliable data on indicators such as PMF, FFD, and HT for LAES and VRFB systems are lacking, as evidenced in Table 2. To address this issue and ensure that all these environmental indicators are still considered in the model, the missing ranges of values for LAES and VRFB systems have been estimated from the other grid energy storage technologies using linear correlations. In this regard, it is noteworthy that very strong linear correlations are observed between the maximum and minimum values of GWP and those of the other three environmental indicators. In fact, the linear fittings yield very high goodness-of-fit coefficients ($R^2 > 0.98$).

Likewise, accurate and comprehensive quantitative information is not available regarding the social indicators considered in this model. In order to address this limitation, these indicators have been evaluated with semantic labels on a five-point categorical scale, ranging from very low to very high. Following the Delphi method, a panel of ten experts in the field was convened to anonymously assign semantic labels to each social indicator for each energy storage technology. The most representative semantic labels of each indicator and technology are presented in Table 3.

Table 3: Model input values of the different indicators for all the grid energy storage technologies considered in the comparative sustainability assessment. The numerical values shown within brackets denote modal values.

Requirement category	Indicator	Unit	Grid energy storage technology					
			PHES	CAES	LAES	NaSB	VRFB	HES
Economic	Levelized cost of storage (LCOS)	\$/kWh	0.02–0.31 (0.10)	0.02–0.28 (0.10)	0.14–0.46	0.29–0.34	0.20–0.56	0.30–0.59 (0.40)
	Power capital cost ($\mathit{\Gamma}_P$)	\$/kW	500–2000	400–1000 (700)	700–3000 (900)	150–3000 (1000)	175–1500 (750)	1000–10000 (3000)
	Capacity capital cost ($\mathit{\Gamma}_C$)	\$/kWh	5–430 (100)	1–270 (35)	100–800 (530)	100–600 (400)	150–1000	45–300
	O&M cost	\$/(kW·yr)	2–52 (15)	2–25 (10)	13.5–30	2–80	4–51	8–48 (17)
	Lifetime (N)	yr	30–60 (50)	20–60 (50)	20–40 (30)	10–15	5–25 (10)	5–30 (20)
Environmental	Global warming potential (GWP)	g CO_2-eq./kWh	23.5–650	27.1–740 (380)	30.2–203	37.9–640	21–433	50.6–1620
	Particulate matter formation (PMF)	g PM_{10}-eq./kWh	0.05–0.85	0.08–1.00	0.08–0.30 [c]	0.08–0.85	0.06–0.6 [c]	0.16–2.18
	Fossil fuels depletion (FFD)	g oil-eq./kWh	5.61–189	7.48–217	8.57–59.22 [c]	9.97–187	5.88–126.63 [c]	17.4–475
	Human toxicity (HT)	g 1,4-DB-eq./kWh	15.7–418 (128)	24.9–475	13.57–122.73 [c]	15.3–417	44–150	35.3–1030
Social	Risk of accidents	Categorical [a]	VL	H	M	L	VL	H
	Population displacement	Categorical [a]	M	H	VL	VL	VL	M
	Visual impact	Categorical [a]	VH	L	M	M	M	L
Technical performance	Technology readiness level (TRL)	1–9	9	7–9	5–7	9	9	4–5
	Round-trip efficiency (η_{rt})	%	65–85 (80)	40–75 (52)	40–70 (50)	70–90	60–85 (65)	20–50 (31)
	Self-discharge rate (SDR)	%/day	0.005–0.02	0.003–0.03	~0	0.05–20	0.20	0.003
	Discharge power rating	MW	100–5000 (2000)	5–320	10–300 (100)	0.05–50 (20)	0.05–25 (5)	0.1–1000 (50)
	Capacity rating	MWh	500–8000	580–2860	10–1000 (400)	1–600 (125)	0.05–100 (10)	100–1000
	Energy density (ρ_E)	kWh/m^3	0.2–2	0.5–20 (6)	20–230 (110)	150–250	10–60 (20)	500–3000
	Geographical constraints	Categorical [b]	Yes	Yes	No	No	No	Yes/No

[a] VH = Very High, H = High, M = Medium, L = Low, VL = Very Low. [b] Yes/No.
[c] Estimated using a linear fitting between this indicator and global warming.

The values in Table 3 reflect the current situation and are expected to remain applicable for the next few years. However, it is important to note that some values may need to be updated and revised over time. Energy storage technologies with low TRL values are the most uncertain and, as such, are more likely to experience changes in the model input values (or distributions) for certain indicators. In these cases, it is reasonable to expect costs to decrease as technologies mature. Following the same reasoning, lifetime is expected to increase for technologies with lower TRL values, while remaining relatively stable for more mature ones. However, if the TRL is too low, overly optimistic estimates could lead to the opposite effects. From an environmental perspective, less mature technologies could see improvements for reasons similar to those outlined for costs, especially in developed countries, where environmental regulations are becoming stricter. Likewise, the technical performance of less mature alternatives may also improve over time, except for the indicator "geographical constraints", which is very unlikely to change. Notably, the "risk of accidents" is the only social indicator included in the model that could vary significantly in the future. There are several reasons for this: increasingly stringent occupational safety regulations, lessons learned from accidents as technologies spread across the globe, and an increase in the TRL, among others. Notwithstanding the above, as long as up-to-date data are used, the model should remain reliable.

3.2. Requirement tree, weights and value functions

Table 4 presents the requirement tree for the model, outlining the various requirements, criteria, and indicators that guide the assessment process. The table is organized into four main requirement categories, each with an associated weight (α_i), and further broken down into ten criteria with their respective weights (β_i). Additionally, the tree includes nineteen indicators, each assigned a specific weight (γ_i). In accordance with the Delphi method, a panel of ten experts in the field was engaged to assign numerical values to the weights of requirements (α_i), criteria (β_i), and indicators (γ_i). The average values for their respective weights are also reported in Table 4.

Table 4: Requirement tree for the model.

α_i	Requirements	β_i	Criteria	γ_i	Indicators
25.6%	Economic	38.0%	Cost of storage	100%	Levelized cost of storage (LCOS)
		21.9%	Short-term financing needs	46.5%	Power capital cost
				53.5%	Capacity capital cost
		20.0%	Long-term financing needs	100%	O&M cost
		20.1%	Risk of obsolescence	100%	Lifetime
28.4%	Environmental	100%	Environmental	29.9%	Global warming potential (GWP)
				23.1%	Particulate matter formation (PMF)
				22.8%	Fossil fuels depletion (FFD)
				24.2%	Human toxicity (HT)
21.9%	Social	100%	Social	40.5%	Risk of accidents
				38.0%	Population displacement
				21.5%	Visual impact
24.1%	Technical	32.9%	Technological maturity	100%	Technology readiness level (TRL)
		30.9%	Efficiency	63.5%	Round-trip efficiency
				36.5%	Self-discharge rate (SDR)
		20.4%	Power and capacity	43.0%	Discharge power rating
				57.0%	Capacity rating
		15.8%	Space requirements	55.0%	Energy density
				45.0%	Geographical constraints

In Table 4, certain criteria, particularly those pertaining to the economic and technical requirement categories, need clarification. The CAPEX criterion represents the short-term financing needs and is assessed through two indicators: power capital cost and capacity capital cost. Conversely, the OPEX criterion addresses long-term financing needs. The lifetime indicator is associated with a risk of premature obsolescence criterion, where technologies with prolonged lifetimes may incur penalties due to a higher probability of becoming obsolete over time. Regarding the technical criteria, the efficiency criterion consists of both round-trip efficiency and self-discharge rate. The size and suitability of each technology for grid energy storage in terms of the power and capacity rating indicators are assessed through the power and capacity criterion. Finally, the space requirements for each technology are evaluated through a criterion that includes the energy density of each technology and the geographical constraints.

A value function was established for each of the indicators under consideration in the comparative sustainability assessment. Table 5 illustrates the parameters used for quantitative value functions. Additionally, Table 6 outlines the satisfaction levels associated with the semantic labels used in the stepped qualitative value functions. All these value functions were developed with varying shapes according to the insights gathered from the aforementioned panel of experts. This variety of shapes reflects the different levels of exigency associated with each indicator.

Table 5: Parameters for the quantitative value functions.

Indicator	Unit	Parameters					Characteristics	
		$P_{i,\max}$	$P_{i,\min}$	a_i	m_i	n_i	Shape	Trend
Levelized cost of storage (LCOS)	\$/kWh	0.02	0.5	4.25	0.5	0.23	S-shaped	Decreasing
Power capital cost (Γ_P)	\$/kW	400	5000	4.8	0.5	2750	S-shaped	Decreasing
Capacity capital cost (Γ_C)	\$/kWh	1	500	4.8	0.5	275	S-shaped	Decreasing
O&M cost	\$/(kW·yr)	2	50	2.7	0.35	20	S-shaped	Decreasing
Lifetime (N)	yr	20	60	2.4	0.8	24	S-shaped	Decreasing
Global warming potential (GWP)	g CO_2-eq./kWh	20	700	2.6	0.2	400	Concave	Decreasing
Particulate matter formation (PMF)	g PM_{10}-eq./kWh	0.05	0.95	2.35	0.16	0.54	Concave	Decreasing
Fossil fuels depletion (FFD)	g oil-eq./kWh	5	200	2.5	0.2	115	Concave	Decreasing
Human toxicity (HT)	g 1,4-DB-eq./kWh	15	400	2.7	0.3	250	Concave	Decreasing
Round trip efficiency (η_{rt})	%	100	0	3	0.5	50	S-shaped	Increasing
Self-discharge rate (SDR)	%/day	0	1	3.5	0.1	0.9	Concave	Decreasing
Discharge power rating	MW	500	0	2.5	0.1	80	S-shaped	Increasing
Capacity rating	MWh	1000	0	2.5	0.1	160	S-shaped	Increasing
Energy density (ρ_E)	kWh/m^3	250	0	0.6	0.5	250	Convex	Increasing

Table 6: Parameters for the qualitative value functions.

Indicator	Semantic label → Satisfaction level [a]	Trend
Risk of accidents	VL → 1, L → 0.8, M → 0.5, H → 0.2, VH → 0	Decreasing
Population displacement	VL → 1, L → 0.8, M → 0.6, H → 0.4, VH → 0	Decreasing
Visual impact	VL → 1, L → 0.8, M → 0.5, H → 0.25, VH → 0	Decreasing
Technology readiness level (TRL)	1–4 → 0, 5 → 0.1, 6 → 0.2, 7 → 0.5, 8 → 0.8, 9 → 1	Increasing
Geographical constraints	No → 1, Yes → 0	Decreasing

[a] VL = Very low, L = Low, M = Medium, H = High, VH = Very High

3.3. Results from the sustainability assessment

The results of the MIVES–Monte Carlo simulation to assess the sustainability index (*SI*) of each grid energy storage technology are presented in Table 7. These results were obtained using the average weights provided in Table 4 and the value functions outlined in Tables 5 and 6. The sustainability index is expressed as a percentage to enhance readability, although it is typically mathematically defined on a scale from 0 to 1. Several parameters are compared across these technologies, starting with the rank, which indicates each technology's relative standing based on the statistical parameters reported in Table 7. The minimum and maximum values represent the lowest and highest sustainability indices observed during the simulations, defining the range of possible outcomes for each technology. The median value reflects the central tendency (50th percentile), showing the point where half of the data points fall below and half above. The mode represents the most frequently occurring *SI* value, while the mean indicates the average sustainability index. The standard deviation provides insights into the variability of the results, with higher values suggesting increased fluctuation, and in turn, greater uncertainty regarding the sustainability performance of each technology.

Table 7: Summary of the MIVES–Monte Carlo simulation results for the sustainability index (*SI*). Even though sustainability indices typically range from 0 to 1, all values are expressed in percentage form to facilitate reading.

Energy storage technology	Parameters							
	Rank	Min. [a]	Median	Max. [b]	Range	Mode	Mean	S. Dev. [c]
PHES	3rd	41.85%	60.04%	78.51%	36.66%	59.39%	60.14%	4.98%
CAES	5th	35.94%	51.86%	71.26%	35.32%	52.48%	52.02%	4.73%
LAES	2nd	49.24%	61.54%	77.19%	27.95%	60.14%	61.86%	4.39%
NaSB	4th	41.70%	54.46%	75.32%	33.61%	54.59%	54.46%	4.81%
VRFB	1st	50.22%	62.81%	80.68%	30.46%	63.45%	63.02%	4.51%
HES	6th	23.30%	33.57%	52.66%	29.35%	32.73%	33.96%	3.93%

[a] Minimum, [b] Maximum, [c] Standard deviation.

Table 7 reveals that VRFB stands out as the most sustainable technology, ranking first with a median sustainability index of 62.81% and a mean value of 63.02%, indicating strong consistency in its performance. LAES and PHES follow closely behind, with median sustainability indices of 61.54% and 60.04%, respectively, and mean values of 61.86% and 60.14%, reflecting minimal deviation between the central tendency and the average performance for these technologies. NaSB and CAES follow at a significant distance, with median values of 54.46% and 51.86%, and corresponding mean values of 54.46% and 52.02%, showing slight variation but still maintaining a relatively close alignment between the median and mean values. Finally, HES ranks last, with the lowest median sustainability index of 33.57% and a mean value of 33.96%, indicating that it is the least sustainable alternative among the technologies assessed.

Fig. 5 illustrates the frequency histograms for the *SI* of the different technologies. These histograms visually represent the information included in Table 7, and are therefore worth interpreting further. On the one hand, the width of the bars on the abscissa axis provides an idea of the range, which serves as a measure of dispersion (once the outliers are discarded) and, therefore, of variability. Technologies with narrower ranges, such as LAES, HES, and VRFB, also exhibit lower standard deviations, another metric for measuring the dispersion or variability of results. Consequently, these technologies demonstrate a more predictable sustainability performance, with *SI* values that are more tightly clustered around the mean. By contrast, PHES, CAES, and NaSB are associated with higher uncertainty, with a wider distribution of *SI* values and higher standard deviations. However, it is important to note that the differences in variability among the technologies are not significant enough to be clearly observed in the frequency histograms of Fig. 5, making Table 7 indispensable. On the other hand, frequency histograms facilitate comparisons between technologies, especially when two alternatives that are not very close in the ranking are pitted against each other. Therefore, the further to the right the histogram is positioned, the better the sustainability performance of the technology. For example, as shown in Fig. 5, the histogram for HES is positioned between *SI* values slightly above 0.2 and those close to 0.5, with most of them clustered between 0.3 and 0.4. The histograms for the other technologies are shifted to the right, with the majority of their *SI* values above 0.4. This indicates that HES is the least sustainable technology. Similarly, from Fig. 5 it is also clear that the histogram for VRFB is positioned between *SI* values close to 0.5 and 0.8, with most of the values concentrated above 0.6. It is therefore obvious that VRFB tends to outperform technologies such as CAES, for instance, which has an *SI* range between 0.4 and 0.7, with its most likely values slightly above 0.5. Other similar comparisons can be drawn from Fig. 5. Nevertheless, such comparisons become more challenging when the technologies exhibit similar performance, as is the case with VRFB and LAES. In other words, Fig. 5 does not clearly reveal which technology is marginally superior in terms of sustainability, because their frequency histograms are quite alike. In such cases, it is essential to refer to the numerical results reported in Table 7.

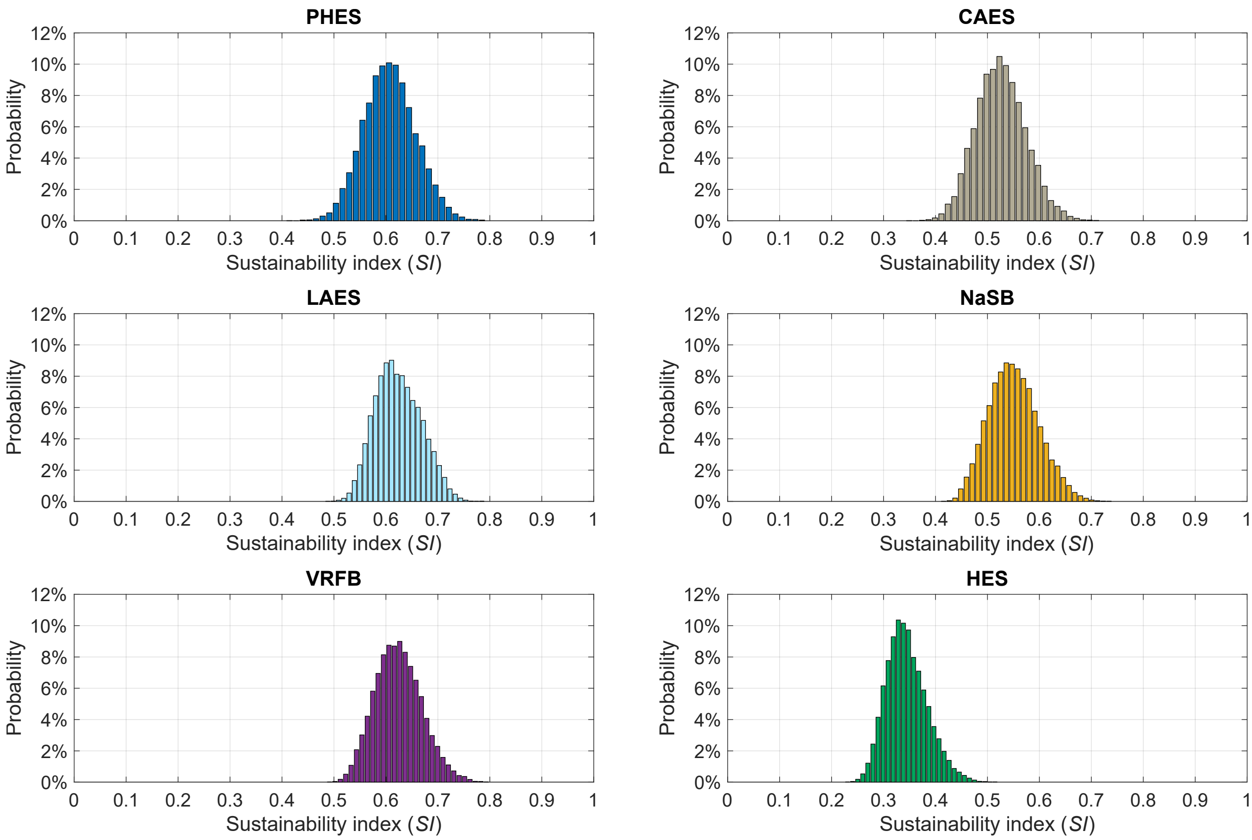

Figure 5: Frequency histograms for the sustainability index (*SI*) of each energy storage technology.

Frequency histograms can also help identify potential asymmetries in the data. However, no significant asymmetries are apparent in the histograms included in Fig. 5. This observation is consistent with the data presented in Table 7, since the mean and median values are always very close to each other. Looking more closely, Table 7 shows that for five of the six technologies, the mean is slightly higher than the median, indicating a minor rightward (or positive) skew. However, given the negligible differences, it is reasonable to conclude that the distributions are nearly symmetrical, as confirmed by Fig. 5.

Fig. 6 illustrates the cumulative probability curves for all the alternatives, facilitating the understanding of the numerical values reported in Table 7. Additionally, it helps to visualize the ranking position of each technology in terms of sustainability performance. Each curve in Fig. 6 represents the cumulative probability of achieving a particular value of the *SI* across the range of potential outcomes. The steepness and position of the curves provide insights into the relative sustainability of each technology. Technologies with curves positioned further to the right exhibit higher sustainability indices, indicating stronger overall performance in sustainability metrics. Conversely, curves positioned to the left reflect lower sustainability indices, suggesting weaker performance.

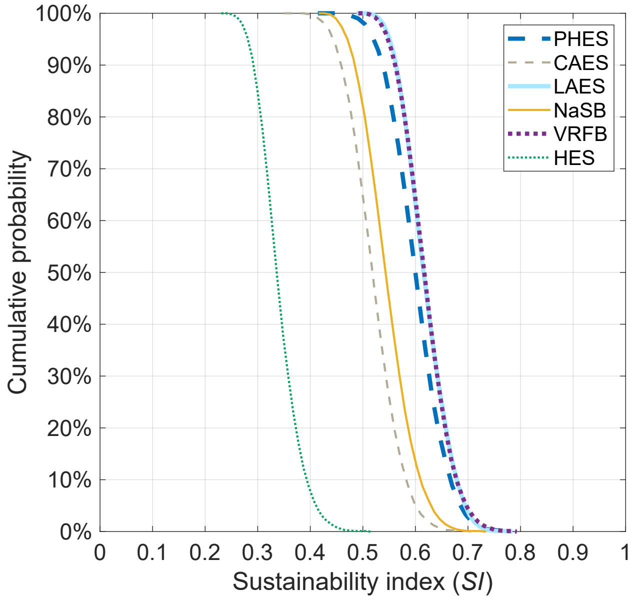

Figure 6: Cumulative probability curves for the sustainability index (*SI*) of each energy storage technology.

As shown in Fig. 6, VFRB generally performs the best, as its curve is positioned furthest to the right, resulting in higher sustainability indices across a wider range of probabilities. LAES and PHES follow at a very close distance, suggesting that these technologies also perform well in terms of sustainability, with their curves overlapping or near VFRB's curve. CAES and NaSB show intermediate sustainability performance, with their curves positioned further to the left toward the middle of the chart, implying lower sustainability indices at comparable cumulative probabilities. Notably, HES ranks the lowest, with its curve positioned the farthest to the left, indicating that it consistently achieves lower sustainability indices compared to the other energy storage technologies.

A more detailed analysis of the results, focusing on each requirement separately, provides additional insights and allows justification of the position of each technology on the basis of the model input values. Fig. 7 displays the cumulative probability curves for the economic, environmental, social, and technical sustainability indices of each alternative.

From an economic point of view, CAES ranks as the top-performing alternative, having obtained the highest modal, mean, minimum, maximum and median values for the economic index. This outcome is reasonable, since CAES presents the most favorable distribution function for the LCOS, closely resembling that of PHES, which ranks as the second-best option in the economic dimension (see Table 3). In fact, the differences between these two technologies mostly lie in the higher short- and long-term financing needs that PHES tends to have (Table 3). The remaining alternatives lag considerably behind these top two options, generally showing higher LCOS values and greater financing needs (Table 3). Although these alternatives have a lower risk of premature obsolescence due to shorter service lives, this advantage does not offset their weaker performance on other economic indicators. Nevertheless, it is noteworthy that some of these technologies, particularly LAES and HES, are not yet mature enough (TRL values ranging from 5 to 7 and from 4 to 5, respectively). Therefore, it is reasonable to expect that as these technologies advance, their increased maturity, together with improvements in process efficiency and experience-driven cost reductions, will drive down overall costs, potentially improving their economic performance over time.

The most significant differences among the six large-scale energy storage technologies occur at the environmental level, with a larger gap between the curves for the best and worst perform-

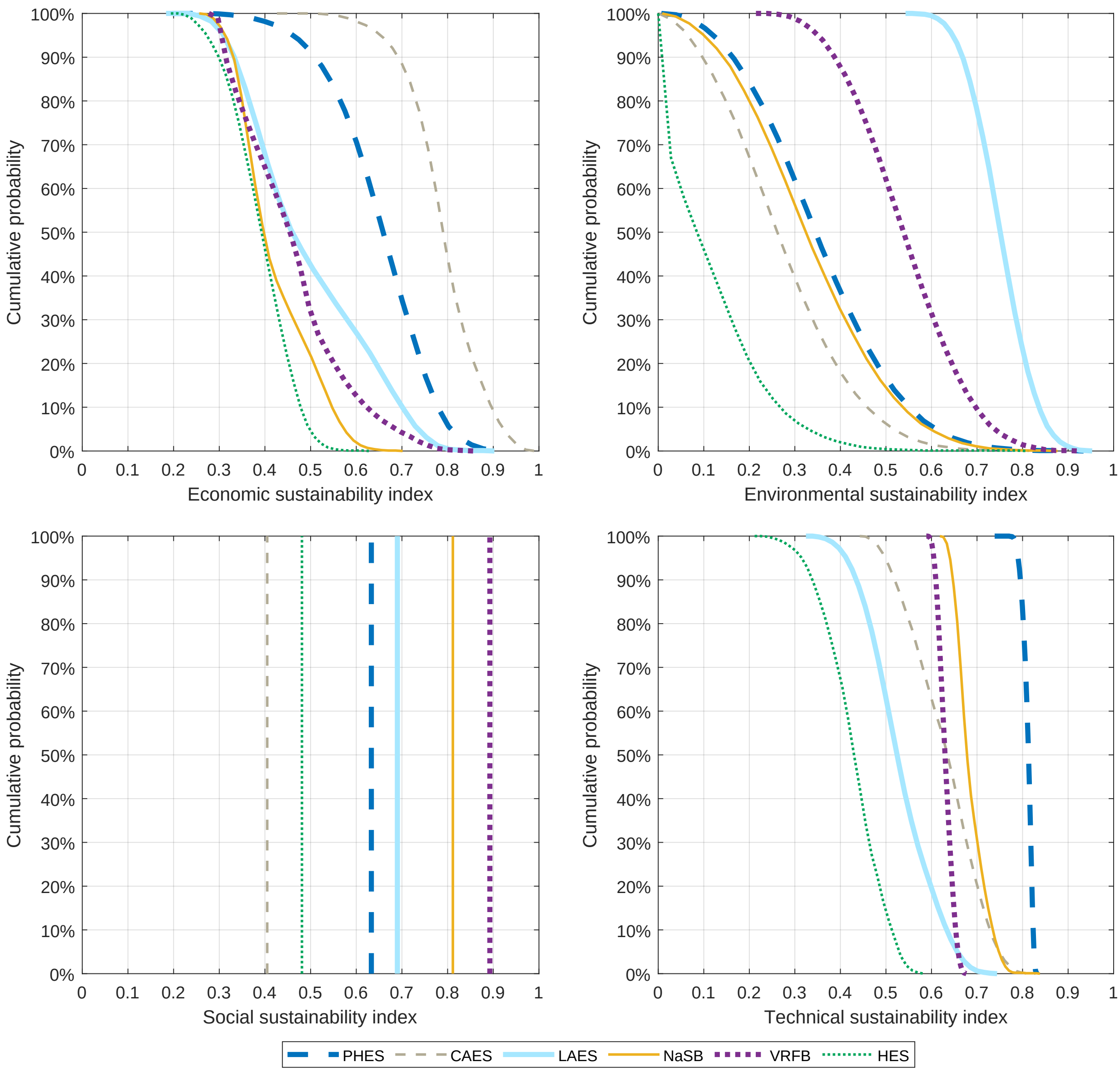

Figure 7: Cumulative probability curves for the partial (economic, environmental, social, and technical) sustainability indices (*SI*) of each energy storage technology.

767 ing options compared to other sustainability dimensions (see Fig. 7). LAES ranks highest in this
768 category, since the maximum values it adopts during simulation across the four environmental in-
769 dicators are considerably lower than those for other technologies, as is clear from Table 3. LAES
770 achieved the best results in terms of minimum, maximum, mean, and median environmental in-
771 dices, as well as the smallest range and standard deviation values, indicating less uncertainty in

its environmental performance relative to other technologies. LAES demonstrates consistently strong environmental results, ranging from acceptable (above 50%) to excellent, while other technologies show a much broader spectrum, ranging from very unsatisfactory to quite satisfactory performances. VRFB ranks as the second-best alternative from an environmental perspective. Although its minimum values are smaller than those of LAES for three of the four environmental indicators, it is handicapped by the fact that it can also adopt considerably higher values (Table 3). Nonetheless, under optimal conditions, its environmental performance can reach an index of over 90%. In contrast, HES ranks last, the same position as in the economic ranking. According to the uniform distribution functions defined in Table 3, this technology currently has the greatest potential to contribute to the four environmental impacts. Consequently, even in the most favorable scenario, its environmental performance is far from the best possible outcomes. The underperformance of HES from an environmental perspective can be partly attributed to its lower round-trip efficiency, resulting in greater energy losses and, consequently, a higher environmental impact per unit of energy delivered back to the grid. The environmental performance of HES also depends on the energy source used for its production, which introduces significant variability in indicators such as GWP. However, it is important to note that HES is still at a relatively low TRL, which means that there is considerable potential for improvements in both efficiency and environmental performance over time.

Regarding the social dimension, the value that each technology adopts for each indicator during simulation remains constant (Table 3). In other words, social indicators were not evaluated probabilistically, resulting in no variability in the outcomes or rankings. VRFB achieved the highest social index (89.25%), followed by NaSB (81.15%). Both technologies are characterized by having a low or very low risk of accidents. Similarly, both present the best possible input value for the indicator of population displacement. The remaining technologies perform considerably worse, with social indices varying from 40.5% (CAES) to 69% (LAES). CAES and HES are arguably the only two technologies with unacceptably low social indices, both below 50%. Their inherent characteristics (high pressures, flammability, potential leakage, among others) contribute to a higher risk of accidents (Table 3). PHES ranks as the fourth option, influenced by its very high visual impact and a medium risk of negative population displacement, which may arise from the need to flood valleys for water reservoirs.

From a technical perspective, PHES is the leading technology, as evidenced by its cumulative probability curve being positioned to the right of the other curves (Fig. 7). Considering this result alongside its second-place ranking in the economic dimension, it can be concluded that PHES is the best option from a techno-economic perspective. Nevertheless, its environmental and social performance prevents it from being deemed the most sustainable technology, although it is well positioned in the overall ranking (Fig. 6). PHES is a mature technology and, consequently, the variability in its technical performance is lower than in other alternatives, with a technical index between 73.82% and 83.36%. Its main drawbacks are low energy density and geographical constraints (Table 3). Conversely, HES currently ranks as the least favorable option. Despite presenting high energy density and negligible self-discharge rates, HES suffers from low TRL and low round-trip efficiency. VRFB and NaSB are quite similar alternatives in terms of technical performance, as indicated by the proximity of their cumulative probability curves in Fig. 7. CAES and LAES show substantial variability in terms of technical performance, with indices ranging from 45.07% to 81.89% for CAES and from 32.42% to 75.28% for LAES. However, both technologies are generally far behind PHES and NaSB. Table 8 summarizes the position that each technology tends to occupy from economic, environmental, social, and technical points of view.

It is important to note that the results presented in this section not only provide an overview of the different levels of sustainability that can be achieved by each energy storage technology. These findings serve as support for decision-makers, since they allow them to know the probability of each technology reaching a level of sustainability above or below a certain threshold. Furthermore, this information is available individually for each sustainability dimension (economic, social, environmental, and technical), which allows decision-makers to select alternatives tailored to the

Table 8: Expected ranking of the grid energy storage technologies in economic, environmental, social, and technical dimensions.

Energy storage technology	Sustainability dimension			
	Economic	Environmental	Social	Technical[a]
PHES	2nd	3rd	4th	1st
CAES	1st	5th	6th	3rd
LAES	3rd	1st	3rd	5th
NaSB	5th	4th	2nd	2nd
VRFB	4th	2nd	1st	3rd
HES	6th	6th	5th	6th

[a] VRFB and CAES share the third position, as in approximately 50% of the iterations, one technology outperforms the other, and vice versa.

specific priorities of a given country or region. These priorities depend on the level of economic development that each country has and are also subject to change over time. However, as with any model based on multi-criteria methods, subjectivity can influence the results, and this study is no exception. Therefore, to reduce the impact of subjectivity, it is advisable to adopt a probabilistic approach and conduct a sensitivity analysis by varying the values of the parameters most affected by subjectivity (further details are provided in Section 3.5).

In summary, no single technology excels across all four dimensions of sustainability. However, some options have demonstrated a more balanced performance, while others exhibit remarkable outcomes in certain dimensions but perform poorly in others. The first group includes VRFB, LAES, and PHES, identified as the three most sustainable alternatives. Each of these technologies ranks first in at least one sustainability dimension and never falls to the last position in any of them. NaSB and CAES are one step below this top tier. NaSB achieves a commendable second position in one dimension but does not rank last in any, which is reflected in its cumulative probability curve being slightly to the left of the top three technologies (see Fig. 6). CAES, on the other hand, ranks fourth, fifth, and sixth in the technical, environmental, and social dimensions, respectively, and these positions do not offset its first-place standing in the economic dimension. HES ranks last in three out of the four dimensions of sustainability, with its best performance being a fifth place in the social dimension. Nevertheless, it should be noted that, despite its lower sustainability performance, HES has a particular advantage over other energy storage systems, which is the versatility of hydrogen. Beyond grid energy storage, hydrogen has a wide range of applications. Due to its higher energy density compared to other energy storage options (see Table 3), hydrogen can be used to power fuel cell vehicles. In industries, hydrogen is essential for ammonia and methanol production, petroleum refining, and steel manufacturing. It also has significant potential in residential heating through hydrogen boilers and blending with natural gas. Additionally, hydrogen is used in electronics, space exploration as a rocket fuel, and as a feedstock in chemical synthesis. These diverse applications of hydrogen may have implications for the overall sustainability of HES that are challenging to quantify, such as the development of associated businesses and ancillary industries. On the other hand, it is important to remember that the results from the sustainability assessment reflect the current state of grid energy storage technologies. Therefore, as these technologies with lower TRL values advance, their sustainability performance is expected to improve over time.

3.4. Comparison with prior research

In this section, the results of this study are compared with those of related studies in the existing literature. However, it should be noted that such comparisons have inherent limitations, since the MIVES method offers distinct advantages over other MCDM approaches, and no previous model has considered as many indicators as the one developed in this study. Nonetheless, partial comparisons remain meaningful for drawing conclusions. Table 9 provides an overview of the key similarities and discrepancies between the results of this study and previous research, along with insights into the main reasons behind such differences.

Ren and Ren [35] developed a model for assessing the sustainability of five energy storage technologies, two of which, PHES and CAES, are also included in this study. They considered a base case and eleven alternative scenarios with modified weights for the indicators. In eight of the twelve scenarios (including the base case), PHES turned out to be a better alternative than

Table 9: Key similarities and differences between the results of this study and prior research.

Source	Common technologies	Coincident results	Differences	Main reasons for the differences
[35]	PHES and CAES	PHES tends to outperform CAES.	CAES has obtained better results.	Absence of social indicators.
[36]	PHES, HES, NaSB and VRFB	• PHES and VRFB tend to be more sustainable than NaSB. • Prioritizing the economic dimension harms HES and favors PHES. • Prioritizing the social dimension boosts VRFB and hinders PHES and HES.	• PHES outperforms VRFB. • HES ranks first.	Excessive differences among the weights of the sustainability dimensions.
[37]	PHES and CAES	PHES tends to outperform CAES.	CAES is closer to PHES.	Absence of social indicators.
[38]	PHES	Not applicable.	Not applicable.	Not applicable.
[39]	PHES, CAES and HES	PHES and CAES have obtained better results than HES.	• CAES outperforms PHES. • CAES, PHES and HES are in close proximity, according to the index used to generate the ranking.	• Inadequate classification of certain indicators. • Use of a scale of 1 to 5 to assess the indicators. • Input values established by only three experts. • Ranking based on a relative index.

CAES, which is consistent with the results obtained in this work. However, the difference in terms of sustainability performance between PHES and CAES is larger in the MIVES model. This is partly due to the consideration of the social dimension of sustainability, where PHES obtained a satisfaction index of 63.30% compared to 40.50% for CAES, the worst alternative from a social perspective. PHES also outperformed CAES in environmental and technical aspects (see Section 3.3 and Fig. 7). If Ren and Ren [35] had included social indicators in their model, PHES would most likely have outperformed CAES in some of the four scenarios in which CAES ranked higher.

Acar et al. [36] proposed a 14-indicator model for assessing the sustainability performance of various energy storage systems considering economic, social, environmental, and technical aspects. In particular, they compared PHES, conventional batteries, high-temperature batteries, flow batteries, and HES. NaSB can be classified as a high-temperature battery, while VRFB is a flow battery. Therefore, they evaluated four of the six alternatives examined in this study. HES was ranked as the most sustainable alternative, followed by PHES, flow batteries (VRFB), conventional batteries, and high-temperature batteries (NaSB), in descending order of sustainability. The results obtained from the MIVES model differ significantly from those of Acar et al. [36], but there are several factors that can explain these discrepancies. One of the main reasons is that the weights introduced in their study do not align with the definitions of sustainable development or the broader concepts of integral and global sustainability. Although slight differences in the weights of sustainability dimensions are acceptable, large disparities contradict the core principles of sustainability. In their study, the technical requirement is almost 4 times more important than the social one and almost 2.5 times more relevant than the economic one. Similarly, the weight defined for the environmental dimension is almost 2.8 times higher than that of the social requirement. Consequently, the weights proposed by Acar et al. [36] do not adequately represent the balance among the diverse aspects of sustainability, and thus, the alternative identified as the best in their model is not necessarily the most sustainable. On the other hand, Acar et al. [36] assessed all indicators for the energy storage systems using semantic labels, each associated with a triangular fuzzy number. Semantic terms are helpful for indicators that cannot be evaluated numerically, or for those that could theoretically be assessed numerically, but for which reliable data are unavailable. However, many of the indicators they considered could have been assessed numerically. Thus, the use of semantic labels is a potential source of inaccuracies. In addition, they included air pollutants such as CO_2, N_2O, CH_4, and fluorinated gases within the environmental dimension. However, a more appropriate approach would have been to consider environmental impact categories (e.g., global warming potential, particulate matter formation, or human toxicity, among others), since a single pollutant can contribute to various environmental impacts. Despite the differences in findings between Acar et al. [36] and this study, there are notable similarities. For instance, they concluded that increasing the weight of the economic requirement reduces the sustainability performance of HES while enhancing that of PHES, which agrees with the results obtained using the MIVES model. Similarly, prioritizing the social requirement improves the performance of flow batteries (VRFB), while PHES and HES decline in the ranking. This aligns with the findings of the MIVES model, which identifies VRFB as the top alternative from a social perspective, with a satisfaction index of 0.89, while PHES and HES fall to the 4th and 5th positions in the ranking, respectively.

Vo et al. [37] also developed a multi-criteria model for assessing the sustainability of energy storage technologies, specifically considering P2G, PHES, and CAES. The focus here is on comparing PHES and CAES, which are also part of this study. They concluded that PHES contributes more to sustainability than CAES, a finding that aligns with the results of this study, where PHES shows higher minimum, maximum, and mean values for the sustainability index than CAES, as previously evidenced in Table 7. Despite this coincidence, it is worth noting that the model proposed by Vo et al. [37] presents several limitations, which may result in misleading conclusions when applied to other energy storage technologies. On the one hand, their model only quantitatively addresses one environmental impact, CO_2-eq. emissions, while the remaining environmental impacts are combined into a single indicator evaluated using semantic labels. On the other hand, a major weakness of their model is the complete absence of social indicators. A model designed to assess sustainability cannot ignore social aspects, even if they are challenging to quantify. At the very least, an evaluation similar to the one performed using the MIVES model should be conducted. Incorporating social indicators into the model developed by Vo et al. [37] would likely increase the difference between PHES and CAES in terms of sustainability performance.

Guo et al. [38] presented a multi-criteria model to assess the sustainability of conventional and underground PHES. As no distinction is made between these two technologies in this work, a direct comparison with their results is not possible. Nevertheless, it is worth clarifying some aspects about the model proposed by Guo et al. [38]. It consists of 16 indicators belonging to the economic, social, and environmental dimensions of sustainability. Technical aspects were overlooked, probably due to the fact that both options (conventional and underground PHES) perform similarly from a technical standpoint. Consequently, the model developed by Guo et al. [38] needs to be completed with technical indicators to be applicable to other alternatives. Otherwise, it may lead to biased results.

Finally, the last model presented in Table 9 corresponds to the one proposed by Albawab et al. [39]. It consists of 17 indicators and was used to compare the following seven energy storage technologies: lead-acid batteries, Li-ion batteries, supercapacitors, HES, CAES, PHES and thermal energy storage. However, the scope of the discussion is limited to HES, CAES, and PHES, as these are the technologies shared by both studies. Their results indicate that CAES is the most sustainable option, followed by PHES and HES, all of which exhibit similar performances. In fact, CAES achieved a degree of utility of 0.990 (with 1 being the highest possible value), while PHES and HES scored 0.910 and 0.881, respectively. It is important to clarify that the degree of utility used by Albawab et al. [39] to rank the technologies is a relative indicator, which is calculated by dividing the result of each option by that of the best-performing one. Consequently, this indicator does not represent the actual level of sustainability of each technology, as one alternative will always have a degree of utility equal to 1, regardless of its contribution to sustainable development. Furthermore, if additional technologies are included in the comparison, the alternative with the degree of utility equal to 1 may change, affecting the degrees of utility of the other options. This issue is avoided with the MIVES method, which, in fact, reveals greater differences in the results among these three technologies. For example, PHES obtained an average *SI* of 0.601, whereas the corresponding *SI* values for CAES and HES were 0.520 and 0.340, respectively. There is a methodological reason to justify this difference. In the MIVES method, the best possible outcome for an indicator (satisfaction equal to 1) does not necessarily correspond to the input value of one of the alternatives. In other words, for a given indicator, none of the alternatives may achieve the optimal value. Therefore, the sustainability indices generated by MIVES are usually lower than those from other methods, as MIVES indicates the actual distance between the sustainability performance of each alternative and the theoretically maximum possible value of sustainability. On the other hand, according to the model presented in this paper, PHES is generally a better option than CAES. In fact, PHES consistently surpasses CAES from a social point of view and usually leads in technical and environmental aspects as well. Only in economic terms does CAES outperform PHES. The model proposed by Albawab et al. [39] presents certain limitations that may explain why CAES obtained a higher contribution to sustainability than PHES. First, they used a scale from 1 (worst possible value) to 5 (best possible value) to evaluate the indicators. The corresponding values for the alternatives were established by three experts. Using these scales, instead of real numerical values with their measurement units (for example, \$/kWh, MWh, g CO_2-eq./kWh, among others), can lead to distortions in the results, as was also discussed for the semantic labels used by Acar et al. [36]. In addition, some of the indicators considered by Albawab et al. [39] were not included in the right dimension of sustainability. For instance, the indicators belonging to the "resource" criterion may be categorized as environmental indicators. Similarly, the indicator "current installed capacity", classified within the social dimension, does not necessarily correlate with positive social outcomes and may, in fact, produce negative effects. Therefore, this indicator should have been disregarded in the model of Albawab et al. [39], as it may lead to biased conclusions. Another possible miscategorization is the "growth rate" indicator, also included in the social dimension. These misclassifications explain why, according to Albawab et al. [39], CAES is more sustainable than PHES, being able to compete not only economically (see Table 3), but also in other dimensions of sustainability. In addition, the environmental assessment is limited to CO_2 and greenhouse gases. All this suggests that the results of Albawab et al. [39] can be further improved and expanded upon, as has been attempted in this work.

3.5. Sensitivity analysis

This section presents a comprehensive sensitivity analysis that evaluates various components of the MIVES–Monte Carlo model. It includes assessments of changes in indicator weights, tree weights, and value functions, as well as an evaluation of indicator impacts using tornado diagrams.

3.5.1. *Sensitivity to changes in average indicator weights*

First of all, the model's sensitivity to variations in its indicators is examined, considering the average weights resulting from the Delphi analysis. As shown in Table 10, the average weights of the four requirement categories are of comparable magnitude, falling within the 20–30% range. Among the criteria, environmental impacts hold the highest weight at 28.4%, followed by social impacts at 21.9%. The remaining criteria exhibit varying weights ranging from 9.7% for the cost of storage to 3.8% for space requirements, while the indicators fluctuate between 9.7% for the LCOS and 1.7% for geographical constraints.

Table 10: Results of the sensitivity analysis considering the average weights resulting from the Delphi analysis.

Criteria	$\alpha_i \beta_i$	Indicators	$\alpha_i \beta_i \gamma_i$
Cost of storage	9.7%	Levelized cost of storage (LCOS)	9.7%
Short-term financing needs	5.6%	Power capital cost	2.6%
		Capacity capital cost	3.0%
Long-term financing needs	5.1%	O&M cost	5.1%
Risk of obsolescence	5.2%	Lifetime	5.2%
Environmental	28.4%	Global warming potential (GWP)	8.5%
		Particulate matter formation (PMF)	6.6%
		Fossil fuels depletion (FFD)	6.5%
		Human toxicity (HT)	6.9%
Social	21.9%	Risk of accidents	8.9%
		Population displacement	8.3%
		Visual impact	4.7%
Technological maturity	7.9%	Technology readiness level (TRL)	7.9%
Efficiency	7.4%	Round-trip efficiency	4.7%
		Self-discharge rate (SDR)	2.7%
Power and capacity	4.9%	Discharge power rating	2.1%
		Capacity rating	2.8%
Space requirements	3.8%	Energy density	2.1%
		Geographical constraints	1.7%

Table 10 reveals that the levelized cost of storage (LCOS) is the most important indicator, accounting for 9.7% of the total weight. Following closely are the risk of accidents at 8.9%, global warming potential at 8.5%, social impacts due to population displacement at 8.3%, and technology readiness level (TRL) at 7.9%. Next, there is a group of indicators with a more limited influence on the model's results, with weights between 6.9% and 4.7%. In descending order of importance these indicators are: human toxicity, particulate matter formation, fossil fuels depletion, lifetime, O&M cost, round-trip efficiency, and visual impact. Finally, the indicators for which the model is less sensitive, with weights equal to or less than 3%, are in descending order of significance: capacity capital cost, capacity rating, self-discharge rate, power capital cost, discharge power rating, energy density, and geographical constraints. However, targeted actions can affect multiple indicators simultaneously. For instance, increasing capital expenditure (related to short-term financing needs) could reduce both O&M costs (related to long-term financing needs) and LCOS. According to Table 10, the model's sensitivity to such an action is 9.2% (9.7% − 2.6% − 3.0% + 5.1%). Larger sensitivity changes can be achieved through other strategies. One such approach involves implementing a value engineering program aimed at reducing investment and O&M costs, thereby decreasing the LCOS. The model's sensitivity to this action is 20.4% (9.7%+2.6%+3.0%+5.1%). This amount could increase to 21.9% (8.5% + 6.6% + 6.9%) if an environmental life cycle assessment (LCA) is conducted, improving emission outcomes. Finally, a more in-depth environmental LCA could increase the sensitivity to 28.4% (8.5% + 6.6% + 6.5% + 6.9%). All these findings are based on the average weights from the opinions of the ten experts consulted in the Delphi analysis.

3.5.2. Sensitivity to changes in indicator values based on tree weights

The second component of this sensitivity analysis involves evaluating the diverse opinions of experts in determining the weights assigned to each requirement (α_i), criterion (β_i) and indicator (γ_i). Table 11 presents the rankings derived from the individual weights assigned by the ten experts consulted in the Delphi analysis [88–93]. Similarly, Fig. 8 includes the cumulative probability curves for the sustainability index (SI) of each energy storage technology according to the eight weight sets presented in Table 11. All participating experts meet the requirements set out by Hallowell and Gambatese [95] for a rigorous application of the Delphi method.

In the first iteration of the Delphi method, the ten experts were thoroughly briefed on the meaning and scope of the requirements, criteria and indicators of the tree. Following this, they

Table 11: *SI* rankings using the different weight sets of the ten experts according to the Delphi method.

Experts	People	PHES	CAES	LAES	NaSB	VRFB	HES
#1, #2, #3	3	1^{st}	5^{th}	2^{nd}	4^{th}	3^{rd}	6^{th}
#4	1	1^{st}	4^{th}	3^{rd}	5^{th}	2^{nd}	6^{th}
#5	1	3^{rd}	4^{th}	1^{st}	5^{th}	2^{nd}	6^{th}
#6	1	4^{th}	5^{th}	2^{nd}	3^{rd}	1^{st}	6^{th}
#7	1	2^{nd}	4^{th}	1^{st}	5^{th}	3^{rd}	6^{th}
#8	1	3^{rd}	5^{th}	1^{st}	4^{th}	2^{nd}	6^{th}
#9	1	3^{rd}	4^{th}	1^{st}	5^{th}	2^{nd}	6^{th}
#10	1	3^{rd}	5^{th}	1^{st}	4^{th}	2^{nd}	6^{th}
Average rank		3^{rd}	5^{th}	1^{st}	4^{th}	2^{nd}	6^{th}

anonymously provided their proposed weights. While the suggestions were largely consistent for some aspects of the requirement tree, notable discrepancies appeared in others. In some instances, these discrepancies can be significant. Nevertheless, the weight sets of most experts led to similar conclusions.

Subjectivity in assigning weights is a potential issue, but it can be managed through multiple rounds of the Delphi method, conducted anonymously. This approach makes it possible to explore the reasons behind differing opinions. The weighting proposals and the rationale used by various experts can thus be shared, with the ultimate aim of reaching a consensus. However, it is also recognized that differing perspectives can persist, and consensus is not guaranteed. For example, experts might assign different weights to the environmental, social, and economic dimensions of sustainability, and such differences may sometimes be irreconcilable. Nevertheless, if consensus is not reached, it remains valuable to consider multiple sets of weights derived by experts and to analyze the resulting differences, rather than relying on a single set of weights, or making decisions based on a single indicator.

In this case, it became clear during the second iteration that achieving a consensus among all the experts was unlikely. Consequently, the model was built using the average weights from the ten experts. Additionally, simulations were conducted for each individual set of weights, and the

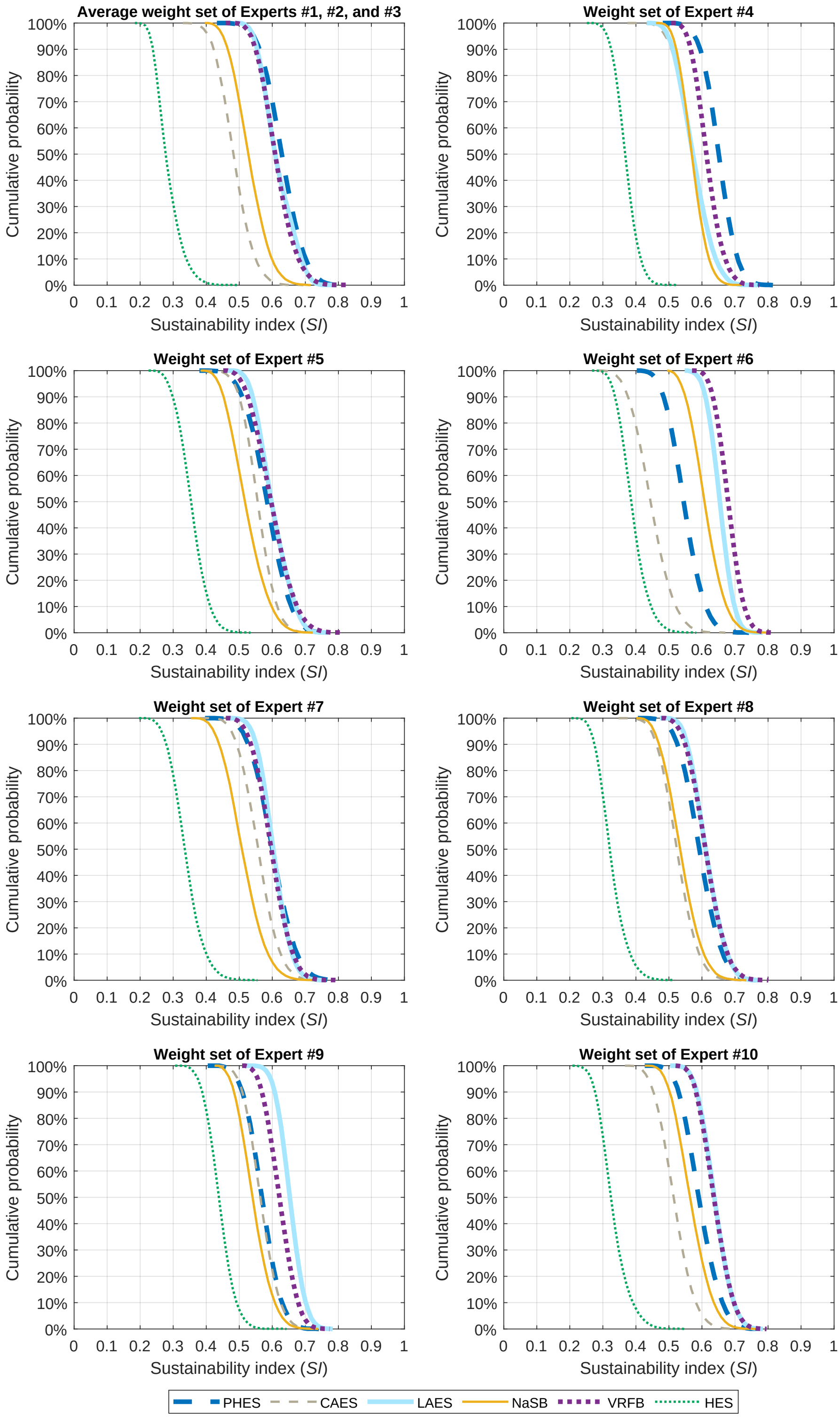

Figure 8: Cumulative probability curves for the sustainability index (*SI*) of each energy storage technology according to the different weight sets.

results were compared as part of this sensitivity analysis.

As shown in Table 11, despite some significant discrepancies in weights assigned by experts, the vast majority of the resulting rankings align closely with the average ranking. These rankings have consistently been established based on the following principles:

- When comparing two options, if their cumulative probability curves do not intersect, the option with the curve further to the right is preferred.
- If the two curves do intersect, the alternative that performs better in a larger number of scenarios is considered the better choice.
- If the two curves intersect around the 50% cumulative probability mark, the option capable of achieving higher sustainability indices is favored.
- When multiple intersections occur, a combination of the latter two criteria is applied.

It is important to note that although the ranking changes when using alternative sets of weights, the overall conclusions in terms of sustainability are essentially the same as those presented in Section 3.3 using average weights. In other words, VRFB, LAES, and PHES remain as the three most sustainable options, although they swap their positions depending on the weight set used. There is only one exception, that is, PHES ranks fourth according to the weights defined by Expert #6. One step below the top three technologies are NaSB and CAES, occupying fourth and fifth positions, in line with the results presented in Section 3.3. There is only one instance where NaSB performs better, reaching third place. Finally, HES is always at the bottom of the ranking, which agrees with the results presented and discussed in Section 3.3. It is therefore possible to conclude that the model provides robust and reliable outcomes, as well as a classification with three clear groups: (i) top-performing technologies (VRFB, LAES, and PHES), (ii) technologies with acceptable performance (NaSB and CAES), and (iii) technology with substantial scope for improvement (HES), regardless of the weight set applied.

3.5.3. Sensitivity to changes in value functions

Reaching consensus on value functions can be even more challenging than agreeing on weights. However, changes in value functions typically have minimal impact on the outcomes of the model. For convenience, a subset of five experts was consulted in this case. These experts agreed to use the value functions with parameters specified in Tables 5 and 6.

In order to analyze the model's sensitivity to changes in value functions, the shapes of both continuous and step functions were adjusted to exhibit a linear trend. The average weights provided by the ten experts were used in this analysis. The results of the MIVES–Monte Carlo simulation with linear value functions are illustrated in Fig. 9.

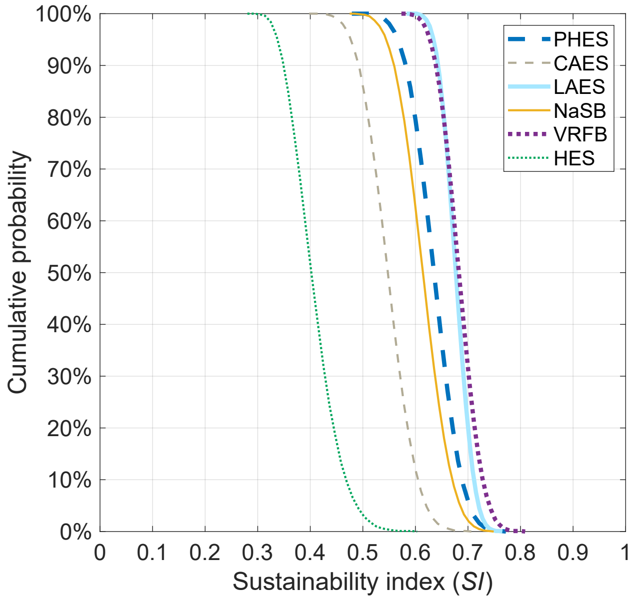

Figure 9: Cumulative probability curves for the sustainability index (*SI*) of each energy storage technology considering linear value functions and using average weights.

Fig. 9 reveals that only two changes in the *SI* ranking occur. Specifically, the top two alternatives exchange their positions, although their cumulative probability curves are very close. Therefore, it is reasonable to expect that minor changes could have resulted in the observed shift in ranking.

3.5.4. *Sensitivity to changes in indicators: tornado diagrams*

In the fourth part of the sensitivity analysis, the base model with average weights was used to evaluate the impact of changes in the indicators on the outcomes of the model. The total sustainability index (*SI*) was calculated using the modal values for triangular distributions and the median values for homogeneous and discrete distributions. Each indicator's value was then varied individually from the minimum to the maximum of its distribution range. Finally, tornado diagrams were generated for each of the six energy storage technologies. These diagrams include the six indicators with the highest influence on the *SI* and are presented collectively in Fig. 10.

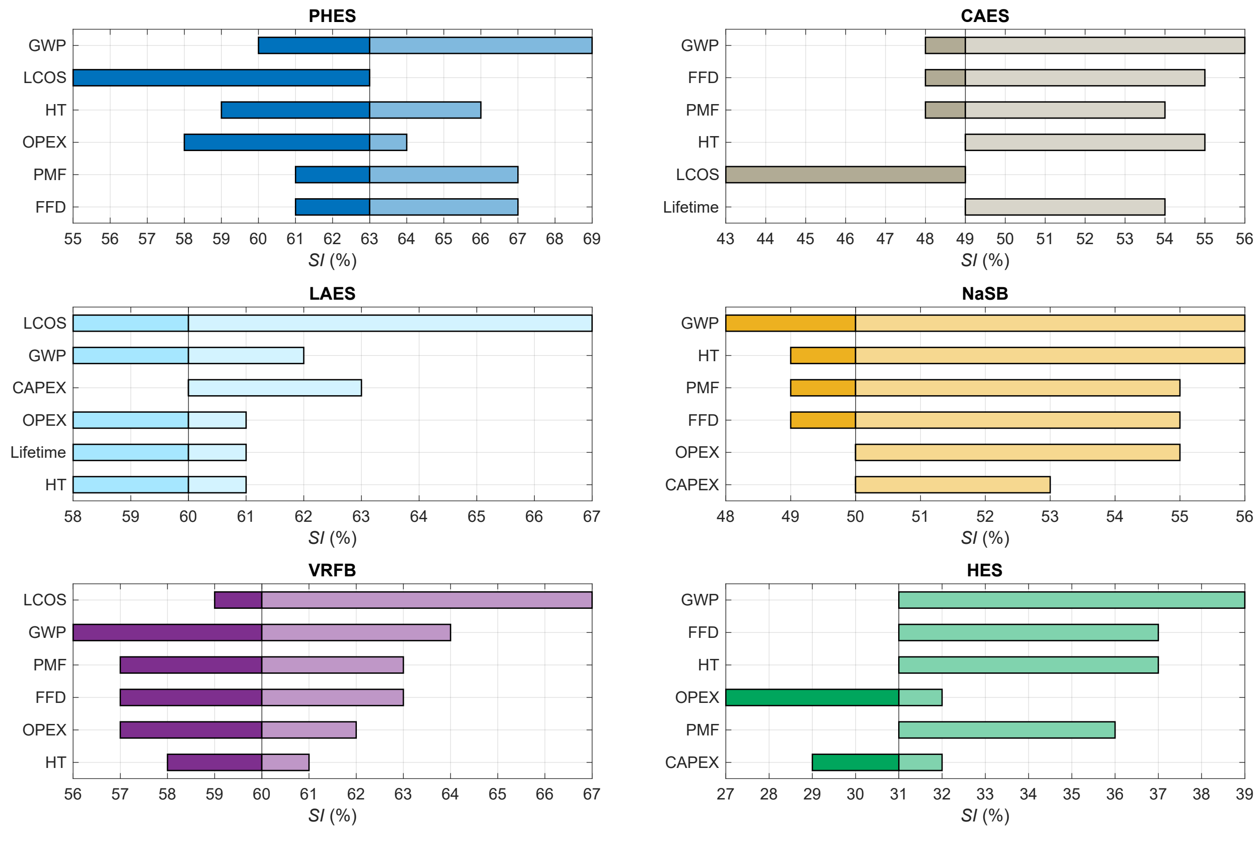

Figure 10: Tornado diagrams of the impact of changes in the indicators on the sustainability index (*SI*) for each energy storage technology.

The tornado diagrams in Fig. 10 reveal that the variations in the *SI* are comparable in magnitude to those observed in the analysis of Section 3.5.1, with no changes exceeding 9%. However, as previously noted, simultaneous changes to multiple indicators can lead to much larger variations, potentially reaching up to 25%.

Some readers might conclude that a probabilistic model is unnecessary if the sensitivity of the model to individual changes in indicator values does not exceed 9%. However, a closer examination of the range and standard deviation of the distribution functions for the sustainability indices reveals substantial variability. This variability becomes evident when looking at the cumulative probability curves. As uncertainty is significant, a probabilistic model is essential to account for the full spectrum of potential scenarios and provide insights into all the possible outcomes.

4. Conclusions, limitations, policy implications and future developments

The MIVES–Monte Carlo method has proven to be an effective approach for assessing the sustainability of large-scale energy storage systems. Unlike traditional approaches, this method makes it possible to incorporate both qualitative and quantitative indicators, consider non-linearities in the assessment, and integrate the analytic hierarchy process (AHP). By using requirement trees and value functions, a comprehensive evaluation of economic, environmental, social, and technical aspects is facilitated, providing a holistic view of sustainability. Indeed, for a comprehensive comparative sustainability assessment of the energy storage systems, it is essential to consider a broad range of indicators. The different sustainability dimensions must be balanced, as technologies that perform well in one area may underperform in others. Moreover, focusing solely on economic or technical aspects may lead to an incomplete and potentially misleading understanding of a technology's overall sustainability. The holistic approach followed in this work provides more accurate and reliable comparisons of the different energy storage technologies, thus supporting decision-making processes in the field of energy storage sustainability. Additionally, the method's flexibility in considering uncertainty through Monte Carlo simulation further enhances its robustness, making it more suitable for sustainability assessment in this context compared to other existing methodologies.

According to the comparative sustainability assessment, there is no energy storage technology with an outstanding performance in all dimensions of sustainability. Nevertheless, the model results suggest that it is possible to clearly distinguish among several groups of technologies according to their most likely contribution to sustainable development. At the first level are vanadium redox flow batteries (VRFB), liquid air energy storage (LAES), and pumped hydroelectric storage (PHES). These technologies consistently occupy the top three positions irrespective of the weight set used (although exchanging positions within them), with only one exception. PHES stands out as the best alternative from a techno-economic standpoint (i.e., the most cost-effective technology), ranking second in economic and first in technical aspects when average weights are applied. LAES achieves the best results from an environmental perspective, while VRFB ranks as the best option from a social point of view. By contrast, hydrogen energy storage (HES), also called power-to-hydrogen-to-power (P2H2P), is identified as the least sustainable option, largely due to its lower technological maturity and higher associated costs. HES consistently ranks last in all partial assessments, with only one exception: it takes fifth place in the social dimension. Positioned at an intermediate level in terms of contribution to sustainable development are sodium-sulfur batteries (NaSB) and compressed air energy storage (CAES), with NaSB showing lower sustainability due to economic and technical aspects, and CAES being penalized by social and environmental concerns. Regarding the results from the sensitivity analysis, the following conclusions have been drawn:

- Reaching an agreement on the weights to be assigned to the different elements of the requirement tree is not always possible. Nevertheless, if these weights are defined according to the concepts of sustainable development and integral sustainability, significant differences in technology rankings are unlikely. In fact, using the weight sets applied in this study, the same conclusion was consistently reached, with only one minor exception. As a result, the technologies were categorized into three groups: the most sustainable (VRFB, LAES, and PHES), those with acceptable performance (NaSB and CAES), and those with room for improvement (HES). The exception involved only a slight shift, with PHES and NaSB swapping places.

- Modifying value functions to adopt linear geometry exerts little influence on the rankings. Only the two best alternatives exchange their positions, both remaining very close to each other in terms of sustainability performance.
- The indicators with the greatest impact on the *SI* were found to be GWP and HT, both of which belong to the environmental dimension. These indicators rank among the top six that most influence *SI* across all energy storage technologies. Following closely are OPEX (economic), PMF (environmental), and FFD (environmental), which also rank among the top six indicators that most influence *SI* for five of the six energy storage technologies.

4.1. Limitations

The model faces several limitations that affect the scope and interpretation of results. A remarkable constraint is the unavailability of more reliable data for some indicators, particularly in the environmental and social dimensions.

For two environmental indicators, namely particulate matter formation (PMF) and fossil fuel depletion (FFD) in technologies such as LAES and VRFB, the absence of complete data required the use of estimations based on linear correlations with other technologies. Despite the very high goodness-of-fit coefficients ($R^2 > 0.98$), this introduces a certain degree of uncertainty into the results, which may not fully represent the specific environmental impacts of these technologies. On the other hand, incorporating additional environmental impact categories, such as those typically used in life cycle assessment methodologies, would also be beneficial to enhance the model's accuracy. Unfortunately, the necessary data for these impact categories was not available.

Similarly, limitations arise in the treatment of social indicators. Due to the challenges in obtaining consistent quantitative data for social impacts, such as population displacement, accident risks, and visual impact, the evaluation relies on expert judgment through the Delphi method. Even though this method helps in filling data gaps, it introduces subjectivity into the assessment, potentially affecting the robustness and accuracy of the social sustainability results. Additionally, the model lacks a social indicator to measure employment generation. This is mainly due to the absence of quantitative data, as well as the difficulty in making qualitative estimations because of the inherent uncertainty of some technologies.

Moreover, the model overlooks certain technical aspects relevant to real-time grid stability, such as rotational inertia and its role in buffering frequency fluctuations. Technologies like PHES, CAES, and LAES, which typically use synchronous electrical machines, can act as spinning reserves and deliver valuable inertial support to the grid, contributing to voltage regulation and helping stabilize frequency during disturbances [96, 97]. Similarly, HES systems, depending on their configuration, can also play a role in maintaining grid frequency and voltage within acceptable limits [97]. In addition to supporting grid stability, some of these technologies can be deployed in black start operations to restore power after a complete grid shutdown. Despite the growing importance of these ancillary services in power systems with increasing penetration of variable renewable energy sources, they were not included in the current set of sustainability indicators and should be considered in future assessments.

Another limitation concerns the subjective nature of the sustainability criteria weighting. Although based on expert input and established methodologies, the weighting of sustainability criteria also remains inherently subjective. Small changes in these weights can slightly alter the ranking of technologies. Therefore, the results should be interpreted with caution, keeping in mind the specific priorities and context of stakeholders. Nevertheless, major changes in the overall conclusions are unlikely, as demonstrated by the sensitivity analysis in Section 3.5. From a methodological standpoint, the model's current version does not include correlations among indicators, since the input data (Table 3) did not provide enough evidence to support establishing such correlations, even through approximate Spearman's correlation coefficients.

In addition, the model does not account for future technological advancements or changes in the regulatory and economic conditions that may significantly alter the sustainability profiles of the storage technologies over time. For instance, technologies like HES, which currently rank lower in sustainability, might improve in the future as innovations reduce costs and increase efficiency, a factor that the current model cannot predict.

Beyond the model's design, a final limitation to consider is the exclusion of certain cutting-edge grid energy storage technologies that could also play a key role in energy transition. This is due to the absence of reliable input data for each indicator of the model for such technologies. However, the model's adaptable design makes it possible to include additional technologies in future assessments, once comprehensive data become available.

4.2. Policy implications

The growing demand for a global energy transition toward a low-carbon economy is driving the need for large-scale adoption of renewable energy sources. This shift requires prioritizing cost-effective electricity generation technologies with low life-cycle environmental and social impacts, such as wind and solar photovoltaic. However, the variability and unpredictability of these renewable technologies pose significant challenges to grid stability and security. Large-scale energy storage systems are therefore essential to facilitate higher renewable energy penetration and avoid the curtailment of excess generation. In this context, energy storage technologies must be carefully selected and deployed to make this energy transition possible. The decisions made will shape the future of the energy sector, impacting not only technical performance but also the broader economic, social, and environmental sustainability of power systems. The comprehensive sustainability assessment model proposed in this research includes a structured approach to evaluating energy storage technologies and provides valuable insights for policy-makers, allowing them to consider not only the technical and economic feasibility of storage options, but also their environmental and social impacts. Based on these findings, some practical policy measures to encourage the adoption of energy storage systems that demonstrate superior environmental and social sustainability outcomes may include offering performance-based government subsidies, facilitating access to state-backed credit lines with more favorable terms than those of private bank loans, or providing tax advantages. Moreover, dedicated funding mechanisms, such as targeted grants or low-interest loans, could be established to support the research, development, and commercialization of less mature yet promising energy storage technologies, thereby fostering a more diverse energy storage landscape. Lastly, incentives for grid operators could be introduced to modernize infrastructure, promoting the development of smart grids that manage renewable energy variability and stabilize the grid, ensuring system flexibility and resilience.

In light of these findings, it is evident that large-scale energy storage systems are critical for integrating the growing installed capacity of variable renewable energy sources and ensuring grid stability. The insights presented in this study offer actionable guidance in formulating energy storage policies that address sustainability across economic, environmental, social, and technical dimensions. This approach is essential for making informed energy policy decisions and driving the balanced energy transition needed to meet global sustainability development goals.

4.3. Future developments

Future developments in this research could focus on expanding the scope of energy storage technologies, allowing for a more comprehensive evaluation of emerging technologies that were not included in the current comparative sustainability assessment. By increasing the number of environmental and social indicators (provided that reliable and complete data for these indicators can be obtained) and by establishing correlations between all the indicators that have been evaluated probabilistically, the robustness of the model would be improved. These steps would lead to a clearer understanding of the interactions between the different sustainability criteria. Another interesting direction for future work could involve using alternative methodologies beyond MIVES, as well as those commonly used in previous studies. With a wide range of MCDM methods available, exploring different approaches could further refine the decision-making process and provide insights into the trade-offs involved in selecting energy storage technologies. Furthermore, the insights gained from this study could be applied to address challenges in generation expansion planning (GEP). This is particularly relevant to scenarios with high penetration of variable renewable energy sources, such as wind and solar power, which are not dispatchable due to their fluctuating nature. As intermittent renewables become more cost-competitive than conventional energy sources, it is essential to factor in the impacts of the storage systems these renewables require. This approach would ensure more accurate and equitable comparisons between conventional and variable renewable energy options.

Nomenclature

Symbols

a	Shape parameter for curve type (convex, concave, linear, or S-shaped)
A	Annual cost [$/yr]
c_e	Unit cost of electricity [$/kWh]
E	Energy [J]
E'	Energy remaining after a storage period [J]
$\dot{E}$	Energy rate [W]
i	Rate of interest
k	Number of indicators
m	Shape parameter for the ordinate at point n
n	Shape parameter for the abscissa at the inflection point
N	Lifetime [yr]
P	Value of any indicator
SI	Sustainability index
R	Residual or recovery value [$/yr]
t	Time [yr]
V	Normalized value of an indicator
W	Work [J]
$\dot{W}$	Power [W]

Greek letters

α	Weight of requirement
β	Weight of criterion
γ	Weight of indicator
Γ	Capital cost
η	Efficiency
ρ	Density

Subscripts

0	Initial condition
f	Final condition
i	i-th element of a sequence, set, or array
C	Capacity
E	Energy
P	Power
re	Reinvestment
rt	Round trip
sd	Self discharge

Abbreviations

AHP	Analytic Hierarchy Process
ARAS	Additive Ratio Assessment

CAES Compressed Air Energy Storage

CAPEX Capital Expenditure

EU European Union

FES Flywheel Energy Storage

FFD Fossil Fuels Depletion

GEP Generation Expansion Planning

GWP Global Warming Potential

HES Hydrogen Energy Storage

HT Human Toxicity

LAES Liquid Air Energy Storage

LCA Life Cycle Assessment

LCOE Levelized Cost of Electricity

LCOS Levelized Cost of Storage

MADA Multi-Attribute Decision Analysis

MAVT Multi-Attribute Value Theory

MCDM Multi-Criteria Decision-Making

MIVES *Modelo Integrado de Valor para una Evaluación Sostenible*

NaSB Sodium–Sulfur Battery

O&M Operation and Maintenance

OPEX Operation and Maintenance Expenditure

P2G Power to Gas

P2H Power to Hydrogen

P2H2P Power to Hydrogen to Power

PHES Pumped Hydroelectric Energy Storage

PMF Particulate Matter Formation

PTES Pumped Thermal Energy Storage

SDR Self-Discharge Rate

SWARA Stepwise Weight Assessment Ratio Analysis

TOPSIS Technique for Order of Preference by Similarity to Ideal Solution

TRL Technology Readiness Level

UK United Kingdom

VRFB Vanadium Redox Flow Battery

WSM Weighted Sum Method

CRediT authorship contribution statement

R. Aguado Molina: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – Original Draft, Writing – Review & Editing. **J.J. Cartelle Barros:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Writing – Original Draft, Writing – Review & Editing. **M.P. de la Cruz López:** Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing – Review & Editing. **M. Lara Coira:** Resources, Supervision, Validation. **A. del Caño Gochi:** Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing – Review & Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

Roque Aguado Molina gratefully acknowledges financial support from Ministerio de Ciencia, Innovación y Universidades (Spain) under the FPU Program (Ref. FPU19/00930).

Funding for open access charge: Universidad de Jaén/CBUA.

Acknowledgments

We would like to thank the Editor and the anonymous reviewers for their constructive and encouraging review, which contributed to improving this paper.

References

[1] H. Chen, T. N. Cong, W. Yang, C. Tan, Y. Li, Y. Ding, Progress in electrical energy storage system: A critical review, Prog. Nat. Sci. 19 (3) (2009) 291–312. doi:10.1016/j.pnsc.2008.07.014.

[2] M. AlShafi, Y. Bicer, Life cycle assessment of compressed air, vanadium redox flow battery, and molten salt systems for renewable energy storage, Energy Reports 7 (2021) 7090–7105. doi:10.1016/j.egyr.2021.09.161.

[3] O. Schmidt, I. Staffell, Monetizing energy storage: A toolkit to assess future cost and value, Oxford University Press, 2023. doi:10.1093/oso/9780192888174.001.0001.

[4] International Energy Agency (IEA), World Energy Outlook 2022, https://www.iea.org/reports/world-energy-outlook-2022, (Accessed: 2024-09-03) (2022).

[5] B. Zakeri, S. Syri, Electrical energy storage systems: A comparative life cycle cost analysis, Renew. Sust. Energ. Rev. 42 (2015) 569–596. doi:10.1016/j.rser.2014.10.011.

[6] International Renewable Energy Agency (IRENA), Electricity Storage and Renewables: Costs and Markets to 2030, https://www.irena.org/publications/2017/Oct/Electricity-storage-and-renewables-costs-and-markets, Accessed: 2024-01-18 (2017).

[7] M. Mann, S. Babinec, V. Putsche, Energy storage grand challenge: Energy storage market report, Tech. rep., National Renewable Energy Laboratory (NREL), Golden, CO (United States) (2020). doi:10.2172/1908714.

[8] International Energy Agency, Global installed energy storage capacity by scenario, 2023 and 2030, https://www.iea.org/data-and-statistics/charts/global-installed-energy-storage-capacity-by-scenario-2023-and-2030, accessed: 2025-03-18 (2024).

[9] European Association for Storage of Energy, Energy Storage Targets 2030 and 2050: Ensuring Europe's Energy Security in a Renewable Energy System, https://ease-storage.eu/wp-content/uploads/2022/06/Energy-Storage-Targets-2030-and-2050-Full-Report.pdf, accessed: 2025-03-18 (2022).

[10] Gobierno de España, Ministerio para la Transición Ecológica y el Reto Demográfico, Estrategia de Almacenamiento Energético, https://www.miteco.gob.es/content/dam/miteco/es/prensa/estrategiaalmacenamiento_tcm30-522655.pdf, accessed: 2025-03-18 (2021).

[11] A. Benato, A. Stoppato, Pumped Thermal Electricity Storage: A technology overview, Therm. Sci. Eng. Prog. 6 (2018) 301–315. doi:10.1016/j.tsep.2018.01.017.

[12] D. Akinyele, R. Rayudu, Review of energy storage technologies for sustainable power networks, Sustain. Energy Technol. Assess. 8 (2014) 74–91. doi:10.1016/j.seta.2014.07.004.

[13] M. C. Argyrou, P. Christodoulides, S. A. Kalogirou, Energy storage for electricity generation and related processes: Technologies appraisal and grid scale applications, Renew. Sustain. Energy Rev. 94 (2018) 804–821. doi:10.1016/j.rser.2018.06.044.

[14] G. F. Frate, L. Ferrari, U. Desideri, Energy storage for grid-scale applications: Technology review and economic feasibility analysis, Renew. Energy 163 (2021) 1754–1772. doi:10.1016/j.renene.2020.10.070.

[15] R. Uría-Martínez, M. M. Johnson, U.s. hydropower market report (2023 edition), Tech. Rep. DOE/EE-2771, U.S. Department of Energy, Oak Ridge National Laboratory, Washington, D.C. (2023).
URL https://www.osti.gov/biblio/2006921

[16] X. Luo, J. Wang, M. Dooner, J. Clarke, C. Krupke, Overview of current development in compressed air energy storage technology, Energy Procedia 62 (2014) 603–611, 6th International Conference on Sustainability in Energy and Buildings, SEB-14. doi:10.1016/j.egypro.2014.12.423.

[17] M. Budt, D. Wolf, R. Span, J. Yan, A review on compressed air energy storage: Basic principles, past milestones and recent developments, Appl. Energy 170 (2016) 250–268. doi:10.1016/j.apenergy.2016.02.108.

[18] O. O'Callaghan, P. Donnellan, Liquid air energy storage systems: A review, Renew. Sustain. Energy Rev. 146 (2021) 111113. doi:10.1016/j.rser.2021.111113.

[19] Y. Ding, L. Tong, P. Zhang, Y. Li, J. Radcliffe, L. Wang, Chapter 9 - liquid air energy storage, in: T. M. Letcher (Ed.), Storing Energy, Elsevier, Oxford, 2016, pp. 167–181. doi:10.1016/B978-0-12-803440-8.00009-9.

[20] E. Borri, A. Tafone, A. Romagnoli, G. Comodi, A review on liquid air energy storage: History, state of the art and recent developments, Renew. Sustain. Energy Rev. 137 (2021) 110572. doi:10.1016/j.rser.2020.110572.

[21] M. Legrand, L. M. Rodríguez-Antón, C. Martinez-Arevalo, F. Gutiérrez-Martín, Integration of liquid air energy storage into the Spanish power grid, Energy 187 (2019) 115965. doi:10.1016/j.energy.2019.115965.

[22] A. M. Salem, A. M. Khaira, A mini-review on liquid air energy storage system hybridization, modelling, and economics: towards carbon neutrality, RSC Adv. 13 (2023) 26380–26391. doi:10.1039/D3RA04506D.

[23] A. M. Rabi, J. Radulovic, J. M. Buick, Comprehensive review of Liquid Air Energy Storage (LAES) technologies, Energies 16 (17) (2023). doi:10.3390/en16176216.

[24] B. Ameel, C. T'Joen, K. De Kerpel, P. De Jaeger, H. Huisseune, M. Van Belleghem, M. De Paepe, Thermodynamic analysis of energy storage with a liquid air Rankine cycle, Appl. Therm. Eng. 52 (1) (2013) 130–140. doi:10.1016/j.applthermaleng.2012.11.037.

[25] P. Krawczyk, Łukasz Szabłowski, S. Karellas, E. Kakaras, K. Badyda, Comparative thermodynamic analysis of compressed air and liquid air energy storage systems, Energy 142 (2018) 46–54. doi:10.1016/j.energy.2017.07.078.

[26] A. Vecchi, K. Knobloch, T. Liang, H. Kildahl, A. Sciacovelli, K. Engelbrecht, Y. Li, Y. Ding, Carnot Battery development: A review on system performance, applications and commercial state-of-the-art, J. Energy Storage 55 (2022) 105782. doi:10.1016/j.est.2022.105782.

[27] A. V. Olympios, J. D. McTigue, P. Farres-Antunez, A. Tafone, A. Romagnoli, Y. Li, Y. Ding, W.-D. Steinmann, L. Wang, H. Chen, et al., Progress and prospects of thermo-mechanical energy storage—a critical review, Progress in Energy 3 (2) (2021) 022001. doi:10.1088/2516-1083/abdbba.

[28] V. Viswanathan, K. Mongird, R. Franks, X. Li, V. Sprenkle, R. Baxter, 2022 Grid Energy Storage Technology Cost and Performance Assessment, Tech. rep., US Department of Energy (2022).

[29] E. D. Spoerke, M. M. Gross, S. J. Percival, L. J. Small, Molten Sodium Batteries, Springer International Publishing, Cham, 2021, pp. 59–84. doi:10.1007/978-3-030-57492-5_3.

[30] A. Risco-Bravo, C. Varela, J. Bartels, E. Zondervan, From green hydrogen to electricity: A review on recent advances, challenges, and opportunities on power-to-hydrogen-to-power systems, Renew. Sustain. Energy Rev. 189 (2024) 113930. doi:10.1016/j.rser.2023.113930.

[31] M. Aneke, M. Wang, Energy storage technologies and real life applications – A state of the art review, Appl. Energy 179 (2016) 350–377. doi:10.1016/j.apenergy.2016.06.097.

[32] X. Luo, J. Wang, M. Dooner, J. Clarke, Overview of current development in electrical energy storage technologies and the application potential in power system operation, Appl. Energy 137 (2015) 511–536. doi:10.1016/j.apenergy.2014.09.081.

[33] M. Yue, H. Lambert, E. Pahon, R. Roche, S. Jemei, D. Hissel, Hydrogen energy systems: A critical review of technologies, applications, trends and challenges, Renew. Sustain. Energy Rev. 146 (2021) 111180. doi:10.1016/j.rser.2021.111180.

[34] N. Schöne, J. Khairallah, B. Heinz, Model-based techno-economic evaluation of power-to-hydrogen-to-power for the electrification of isolated African off-grid communities, Energy Sustain. Dev. 70 (2022) 592–608. doi:10.1016/j.esd.2022.08.020.

[35] J. Ren, X. Ren, Sustainability ranking of energy storage technologies under uncertainties, J. Clean. Prod. 170 (2018) 1387–1398. doi:10.1016/j.jclepro.2017.09.229.

[36] C. Acar, A. Beskese, G. T. Temur, A novel multicriteria sustainability investigation of energy storage systems, Int. J. Energy Res. 43 (12) (2019) 6419–6441. doi:10.1002/er.4459.

[37] T. T. Q. Vo, A. Xia, F. Rogan, D. M. Wall, J. D. Murphy, Sustainability assessment of large-scale storage technologies for surplus electricity using group multi-criteria decision analysis, Clean Technol. Environ. Policy 19 (3) (2019) 689–703. doi:10.1007/s10098-016-1250-8.

[38] Z. Guo, S. Ge, X. Yao, H. Li, X. Li, Life cycle sustainability assessment of pumped hydro energy storage, Int. J. Energy Res. 44 (1) (2020) 192–204. doi:10.1002/er.4890.

[39] M. Albawab, C. Ghenai, M. Bettayeb, I. Janajreh, Sustainability performance index for ranking energy storage technologies using multi-criteria decision-making model and hybrid computational method, J. Energy Storage 32 (2020) 101820. doi:10.1016/j.est.2020.101820.

[40] International Energy Agency (IEA), Batteries and Secure Energy Transitions, https://www.iea.org/reports/batteries-and-secure-energy-transitions (2024).

[41] M. P. de la Cruz, A. Castro, A. del Caño, D. Gómez, M. Lara, J. J. Cartelle, Comprehensive methods for dealing with uncertainty in assessing sustainability part 1: The MIVES–Monte Carlo method, in: Soft computing applications for renewable energy and energy efficiency, IGI Global, 2015, pp. 69–106.

[42] A. Aguado, A. del Caño, M. P. de la Cruz, D. Gómez, A. Josa, Sustainability assessment of concrete structures within the spanish structural concrete code, J. Constr. Eng. Manag. 138 (2) (2012) 268–276.

[43] O. Pons, M. M. Casanovas-Rubio, J. Armengou, A. de la Fuente, Sustainability-driven decision-making model:

Case study of fiber-reinforced concrete foundation piles, J. Constr. Eng. Manag. 147 (10) (2021) 04021116. doi:10.1061/(ASCE)CO.1943-7862.0002073.

[44] A. de la Fuente, M. M. Casanovas-Rubio, O. Pons, J. Armengou, Sustainability of column-supported RC slabs: fiber reinforcement as an alternative, J. Constr. Eng. Manag. 145 (7) (2019) 04019042.

[45] I. Josa, O. Pons, A. de la Fuente, A. Aguado, Multi-criteria decision-making model to assess the sustainability of girders and trusses: Case study for roofs of sports halls, J. Clean. Prod. 249 (2020) 119312. doi:10.1016/j.jclepro.2019.119312.

[46] J. Cuadrado, M. Zubizarreta, B. Pelaz, I. Marcos, Methodology to assess the environmental sustainability of timber structures, Constr. Build. Mater. 86 (2015) 149–158. doi:10.1016/j.conbuildmat.2015.03.109.

[47] M. Zubizarreta, J. Cuadrado, A. Orbe, H. García, Modeling the environmental sustainability of timber structures: A case study, Environ. Impact Assess. Rev. 78 (2019) 106286. doi:10.1016/j.eiar.2019.106286.

[48] M. M. Casanovas-Rubio, P. Pujadas, F. Pardo-Bosch, A. Blanco, A. Aguado, Sustainability assessment of trenches including the new eco-trench: A multi-criteria decision-making tool, J. Clean. Prod. 238 (2019) 117957. doi:10.1016/j.jclepro.2019.117957.

[49] J.-T. San-José Lombera, J. Cuadrado Rojo, Industrial building design stage based on a system approach to their environmental sustainability, Constr. Build. Mater. 24 (4) (2010) 438–447. doi:10.1016/j.conbuildmat.2009.10.019.

[50] J. Cuadrado, E. Rojí, J. T. S. José, J. P. Reyes, Sustainability index for industrial buildings, Proc. Inst. Civ. Eng.: Struct. Build. 165 (5) (2012) 245–253.

[51] O. Pons, A. Aguado, Integrated value model for sustainable assessment applied to technologies used to build schools in Catalonia, Spain, Build. Environ. 53 (2012) 49–58. doi:10.1016/j.buildenv.2012.01.007.

[52] M. M. Casanovas-Rubio, J. Armengou, Decision-making tool for the optimal selection of a domestic water-heating system considering economic, environmental and social criteria: Application to Barcelona (Spain), Renew. Sustain. Energy Rev. 91 (2018) 741–753. doi:10.1016/j.rser.2018.04.040.

[53] M. Karmele Herranz-Pascual, J. Luis Eguiguren-Garcia, R. Proy-Rodriguez, J. Cuadrado-Rojo, New tools to support decision making in urban planning. Model of sustainability assessment of municipal action plans, Dyna 88 (4) (2013) 462–472.

[54] J. J. Cartelle Barros, M. Lara Coira, M. P. de la Cruz López, A. del Caño Gochi, Assessing the global sustainability of different electricity generation systems, Energy 89 (2015) 473–489. doi:10.1016/j.energy.2015.05.110.

[55] J. J. Cartelle Barros, M. Lara Coira, M. P. de la Cruz López, A. del Caño Gochi, Probabilistic life-cycle cost anal-

ysis for renewable and non-renewable power plants, Energy 112 (2016) 774–787. doi:10.1016/j.energy.2016.06.098.

[56] J. J. Cartelle Barros, M. Lara Coira, M. P. de la Cruz López, A. del Caño Gochi, I. Soares, Probabilistic multicriteria environmental assessment of power plants: A global approach, Appl. Energy 260 (2020) 114344. doi:10.1016/j.apenergy.2019.114344.

[57] V. Jülch, Comparison of electricity storage options using levelized cost of storage (LCOS) method, Appl. Energy 183 (2016) 1594–1606. doi:10.1016/j.apenergy.2016.08.165.

[58] G. Assefa, B. Frostell, Social sustainability and social acceptance in technology assessment: A case study of energy technologies, Technol. Soc. 29 (1) (2007) 63–78. doi:10.1016/j.techsoc.2006.10.007.

[59] A. Kumar, A. R. Singh, Y. Deng, X. He, P. Kumar, R. C. Bansal, Integrated assessment of a sustainable microgrid for a remote village in hilly region, Energy Convers. Manag. 180 (2019) 442–472. doi:10.1016/j.enconman.2018.10.084.

[60] H. Afshari, S. Agnihotri, C. Searcy, M. Jaber, Social sustainability indicators: A comprehensive review with application in the energy sector, Sustain. Prod. Consum. 31 (2022) 263–286. doi:10.1016/j.spc.2022.02.018.

[61] M. Héder, From NASA to EU: the evolution of the TRL scale in Public Sector Innovation, The Innovation Journal 22 (2) (2017) 1–23.

[62] A. G. Ter-Gazarian, Energy storage as a structural unit of a power system, in: Energy Storage for Power Systems, The Institution of Engineering and Technology, 2024, pp. 31–42. doi:10.1049/PBPO063E_ch2.

[63] M. Obi, S. Jensen, J. B. Ferris, R. B. Bass, Calculation of levelized costs of electricity for various electrical energy storage systems, Renew. Sust. Energy Rev. 67 (2017) 908–920. doi:10.1016/j.rser.2016.09.043.

[64] S. Kapila, A. Oni, E. Gemechu, A. Kumar, Development of net energy ratios and life cycle greenhouse gas emissions of large-scale mechanical energy storage systems, Energy 170 (2019) 592–603. doi:10.1016/j.energy.2018.12.183.

[65] G. Caralis, T. Christakopoulos, S. Karellas, Z. Gao, Analysis of energy storage systems to exploit wind energy curtailment in Crete, Renew. Sustain. Energy Rev. 103 (2019) 122–139. doi:10.1016/j.rser.2018.12.017.

[66] S. Georgiou, N. Shah, C. N. Markides, A thermo-economic analysis and comparison of pumped-thermal and liquid-air electricity storage systems, Appl. Energy 226 (2018) 1119–1133. doi:10.1016/j.apenergy.2018.04.128.

[67] M. M. Rahman, A. O. Oni, E. Gemechu, A. Kumar, Assessment of energy storage technologies: A review, Energy Convers. Manag. 223 (2020) 113295. doi:10.1016/j.enconman.2020.113295.

[68] P. Marocco, D. Ferrero, M. Gandiglio, M. Ortiz, K. Sundseth, A. Lanzini, M. Santarelli, A study of the techno-economic feasibility of H_2-based energy storage systems in remote areas, Energy Convers. Manag. 211 (2020) 112768. doi:10.1016/j.enconman.2020.112768.

[69] D. Ferrero, M. Gamba, A. Lanzini, M. Santarelli, Power-to-Gas hydrogen: Techno-economic assessment of processes towards a multi-purpose energy carrier, Energy Procedia 101 (2016) 50–57. doi:10.1016/j.egypro.2016.11.007.

[70] A. Gallo, J. Simões-Moreira, H. Costa, M. Santos, E. Moutinho dos Santos, Energy storage in the energy transition context: A technology review, Renew. Sustain. Energy Rev. 65 (2016) 800–822. doi:10.1016/j.rser.2016.07.028.

[71] S. Rehman, L. M. Al-Hadhrami, M. M. Alam, Pumped hydro energy storage system: A technological review, Renew. Sustain. Energy Rev. 44 (2015) 586–598. doi:10.1016/j.rser.2014.12.040.

[72] R. Amirante, E. Cassone, E. Distaso, P. Tamburrano, Overview on recent developments in energy storage: Mechanical, electrochemical and hydrogen technologies, Energy Convers. Manag. 132 (2017) 372–387. doi:10.1016/j.enconman.2016.11.046.

[73] O. Burian, P. Dančová, Compressed air energy storage (CAES) and liquid air energy storage (LAES) technologies—A comparison review of technology possibilities, Processes 11 (11) (2023). doi:10.3390/pr11113061.

[74] E. Bazdar, M. Sameti, F. Nasiri, F. Haghighat, Compressed air energy storage in integrated energy systems: A review, Renew. Sustain. Energy Rev. 167 (2022) 112701. doi:10.1016/j.rser.2022.112701.

[75] G. Venkataramani, P. Parankusam, V. Ramalingam, J. Wang, A review on compressed air energy storage – A pathway for smart grid and polygeneration, Renew. Sustain. Energy Rev. 62 (2016) 895–907. doi:10.1016/j.rser.2016.05.002.

[76] L. Oliveira, M. Messagie, J. Mertens, H. Laget, T. Coosemans, J. van Mierlo, Environmental performance of electricity storage systems for grid applications, a life cycle approach, Energy Convers. Manag. 101 (2015) 326–335. doi:10.1016/j.enconman.2015.05.063.

[77] T. R. Simon, D. Inman, R. Hanes, G. Avery, D. Hettinger, G. Heath, Life cycle assessment of closed-loop pumped storage hydropower in the united states, Environ. Sci. Technol. 57 (33) (2023) 12251–12258. doi:10.1021/acs.est.2c09189.

[78] A. Abdon, X. Zhang, D. Parra, M. K. Patel, C. Bauer, J. Worlitschek, Techno-economic and environmental assessment of stationary electricity storage technologies for different time scales, Energy 139 (2017) 1173–1187. doi:10.1016/j.energy.2017.07.097.

[79] P. Denholm, G. L. Kulcinski, Life cycle energy requirements and greenhouse gas emissions from large scale

energy storage systems, Energy Convers. Manag. 45 (13) (2004) 2153–2172. doi:10.1016/j.enconman.2003.10.014.

[80] E. A. Bouman, M. M. Øberg, E. G. Hertwich, Environmental impacts of balancing offshore wind power with compressed air energy storage (CAES), Energy 95 (2016) 91–98. doi:10.1016/j.energy.2015.11.041.

[81] C. Jones, P. Gilbert, L. Stamford, Assessing the climate change mitigation potential of stationary energy storage for electricity grid services, Environ. Sci. Technol. 54 (1) (2020) 67–75. doi:10.1021/acs.est.9b06231.

[82] S. Weber, J. F. Peters, M. Baumann, M. Weil, Life Cycle Assessment of a Vanadium Redox Flow Battery, Environ. Sci. Technol. 52 (18) (2018) 10864–10873. doi:10.1021/acs.est.8b02073.

[83] A. Escamilla, D. Sánchez, L. García-Rodríguez, Assessment of power-to-power renewable energy storage based on the smart integration of hydrogen and micro gas turbine technologies, Int. J. Hydrog. Energy 47 (40) (2022) 17505–17525. doi:10.1016/j.ijhydene.2022.03.238.

[84] H. Zhao, Q. Wu, S. Hu, H. Xu, C. N. Rasmussen, Review of energy storage system for wind power integration support, Appl. Energy 137 (2015) 545–553. doi:10.1016/j.apenergy.2014.04.103.

[85] O. Dumont, G. F. Frate, A. Pillai, S. Lecompte, M. de Paepe, V. Lemort, Carnot battery technology: A state-of-the-art review, J. Energy Storage 32 (2020) 101756. doi:10.1016/j.est.2020.101756.

[86] S. Hong, C. J. Bradshaw, B. W. Brook, Evaluating options for sustainable energy mixes in South Korea using scenario analysis, Energy 52 (2013) 237–244. doi:10.1016/j.energy.2013.02.010.

[87] M. Dombi, I. Kuti, P. Balogh, Sustainability assessment of renewable power and heat generation technologies, Energy Policy 67 (2014) 264–271. doi:10.1016/j.enpol.2013.12.032.

[88] N. Dalkey, O. Helmer, An experimental application of the Delphi method to the use of experts, Manag. Sci. 9 (3) (1963) 351–515. doi:10.1287/mnsc.9.3.458.

[89] J. Pill, The Delphi method: Substance, context, a critique and an annotated bibliography, Socio-Econ. Plan. Sci. 5 (1) (1971) 57–71. doi:10.1016/0038-0121(71)90041-30.

[90] A. Van de Ven, A. Delbecq, The effectiveness of nominal, Delphi, and interacting group decision making processes, Acad. Manag. J. 17 (4) (1974) 605–621. doi:10.5465/255641.

[91] H. Lindstone, M. Turoff (Eds.), The Delphi Method, Addison Wesley, 1975.

[92] R. Loo, The Delphi method: a powerful tool for strategic management, Polic. Int. J. 25 (4) (2002) 762–769. doi:10.1108/13639510210450677.

[93] C. Okoli, S. Pawlowski, The Delphi method as a research tool: an example, design considerations and applications, Inf. Manag. 42 (1) (2004) 15–29. doi:10.1016/j.im.2003.11.002.

[94] T. L. Saaty, Fundamentals of decision making and priority theory with the analytic hierarchy process, RWS

publications, 1994.

[95] M. Hallowell, J. Gambatese, Qualitative Research: Application of the Delphi Method to CEM Research, J. Constr. Eng. Manag. 136 (1) (2010). doi:10.1061/(ASCE)CO.1943-7862.0000137.

[96] M. Kheshti, X. Zhao, T. Liang, B. Nie, Y. Ding, D. Greaves, Liquid air energy storage for ancillary services in an integrated hybrid renewable system, Renew. Energy 199 (2022) 298–307. doi:10.1016/j.renene.2022.09.010.

[97] F. Ahmed, D. Al Kez, S. McLoone, R. J. Best, C. Cameron, A. Foley, Dynamic grid stability in low carbon power systems with minimum inertia, Renew. Energy 210 (2023) 486–506. doi:10.1016/j.renene.2023.03.082.