Technological Forecasting & Social Change 209 (2024) 123816

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

A holistic approach to assessing reliability in green hydrogen supply chains using mixed methods

Sofía De-León Almaraz [a,*], Tchougoune Moustapha Mai [b], Iris Rocio Melendez [a], M.K. Loganathan [c,d], Catherine Azzaro-Pantel [e]

[a] Institute of Operations and Decision Sciences, Department of Supply Chain Management, Corvinus University of Budapest, Hungary
[b] UMR CNRS 6134 Renewable Energy Laboratory, Scientific Centre Georges Peri, University of Corsica, Ajaccio, France
[c] School of Engineering and Technology, Suresh Gyan Vihar University, Jaipur, India
[d] The Assam Kaziranga University, Jorhat, Rowmarikhuti, Assam 785006, India
[e] Laboratoire de Génie Chimique, Université de Toulouse, CNRS, INPT, UPS, Toulouse, France

ARTICLE INFO

Keywords:
Hydrogen supply chain
Reliability assessment
Mixed methods research
Multi-attribute utility theory
Horizon scanning
Technological expectations

ABSTRACT

Estimating the reliability of future energy supply chains is a vital yet complex task driven by environmental and energy security concerns in the context of the ongoing energy transition. This transition necessitates the integration of new technologies and systems into interconnected networks or supply chains. In this context, hydrogen plays a crucial role in the transition to green energy, as it is anticipated a surge in the establishment of “green” hydrogen supply chains (HSC), necessitating the assurance of reliability in meeting international roadmap targets. Technological reliability is typically evaluated by applying quantitative methods to current technologies. For future HSCs, the reliability assessment challenge is related to their prospective nature, with additional uncertainty due to the technologies' interdependencies. When stakeholders rely solely on technology readiness levels, essential aspects of the supply chain are not considered. This work introduces a novel methodology to assess the technological and organizational reliability of future HSCs, contributing to the literature on hydrogen reliability and strategic foresight. It also offers macro-level reliability projections for green HSCs by 2030, integrating input from energy experts and providing valuable insights for the scientific community, academia, and professionals. The proposed methodology's novelty lies in its ability to integrate various nodes of prospective HSCs. The study employs mixed methods, incorporating quantitative (multi-attribute utility theory) and qualitative approaches (horizon scanning). Variables such as capacity, flexibility, infrastructure vulnerability, and consequences of disruption are considered to quantify reliability, with twenty-four metrics included. Data collection employs the perspective of 2030 through a participatory study based on surveys and interviews, drawing insights from twenty-nine international experts associated with various HSCs-related technologies. The methodology is applied to a case study for a green HSC involving solar/wind energy, electrolysis, transportation, storage, and refueling stations. This paper presents the quantitative results, projecting moderate reliability for green HSCs by 2030. Solar HSCs have been considered slightly more reliable than wind HSCs. The interdependence of electrolysis technology and several aspects related to hydrogen transportation are perceived as vital risks affecting the reliability of green HSCs. Having a constant hydrogen supply is seen as a more significant challenge than HSC's response to unexpected interruptions. The research found specific disparities in expert opinions that enriched the data collection process with complementary viewpoints, benefiting from the former's heterogeneous profiles.

Abbreviations

BBN	Bayesian (Belief) Network
BT	Bow-Tie (Analysis/Diagram)
CH2	compressed hydrogen storage system
CRADIS	compromise ranking of alternatives from distance to ideal solution
CRITIC	combining criteria interaction through inter-criteria

---

* Corresponding author.
E-mail addresses: de.sofia@uni-corvinus.hu (S. De-León Almaraz), moustapha-mai_m@univ-corse.fr (T. Moustapha Mai), iris.melendez@uni-corvinus.hu (I.R. Melendez), mk.loganathan@iitdalumni.com (M.K. Loganathan), catherine.azzaropantel@toulouse-inp.fr (C. Azzaro-Pantel).

https://doi.org/10.1016/j.techfore.2024.123816
Received 11 December 2023; Received in revised form 4 September 2024; Accepted 5 October 2024

	correlation
EDNS	Expected Demand Not Supplied
ESD	Event Sequence Diagrams
ETA	Event Tree Analysis
FEM	Finite Element Model Kriging method
FMEA	Failure Mode and Effect Analysis
FMECA	Failure Mode Effect and Criticality Analysis
FMME	failure modes and maintenance events
FORM	First Order Reliability Method
FST	Fuzzy Set Theory
FTA	Fault Tree Analysis
GOs	General Objectives
HAZID	Hazard Identification
HAZOP	Hazard and Operability Studies
HL	Hasofer and Lind
HRS	hydrogen refueling station
HS	Horizon Scanning
HSC	hydrogen supply chains
HyRAM	hydrogen risk assessment models
ISRM	Importance Sampling Reduction Method
LH2	liquid hydrogen storage system
LOHC	Liquid Organic Hydrogen Carriers
LOHP	Loss of Hydrogen Probability
LOLP	Loss of Load Probability
LPSP	Loss of Power Supply Probability
MA	Markov Analysis
MADM	Multi Attribute Decision Making
MAUT	multi-attribute utility theory
MCDA	Multi Criteria Decision Analysis
MCS	Monte Carlo Simulation
NG	Natural Gas
PEM	Proton Exchange membrane
PoE	Probability of Exceedance
PoF	Probability of Failure
PtG	Power to Gas
PV	Photovoltaic energy
RA	reliability analyses
RBD	Reliability Block Diagram
RI	Reliability Index
SMR	Steam Methane Reforming
SORM	Second Order Reliability Method
SRSM	Stochastic Response Surface Method
SWARA	Stepwise Weight Assessment Ratio Analysis
SWIFT	Structure What-if Technique
SWOT	Strengths, Weaknesses, Opportunities, and Threats
TRL	technological readiness level
UGF	Universal Generating Function
WASPAS	Weighted Aggregated Sum Product Assessment

Sets, parameters, and variables

c	*concepts*
e	*experts opinions*
i	*metrics*
j	*subsystems*
n	*attributes*
r	*General Objectives (Adequacy and Security)*

Variables

$\overline{w_{c,j}}$	*average importance rate for concepts (c) for technology (j)*
$\overline{w_{c,j^*}}$	*average importance rate for concepts (c) included in objective (r) for technology (j)*
$\overline{w_{i,j}}$	*average importance rate for metric (i) for technology (j)*
$\overline{w_{i,j^*}}$	*average importance rate for metric (i) included in concept (c) for technology (j)*
$MaxLikertValue$	*highest allowed value in the survey: 5*
$N_{e,j}$	*total number of experts (e) evaluating technology (j)*
$u_{c,j}$	*utility per concept*
$u_{i,j}$	*normalized utility for metric (i) for technology (j)*
$u_{i,j,e}$	*reliability rate for metric (i) for technology (j) given by expert (e)*
$U_{r,j}$	*utility per general objective*
$w_{c,j}$	*normalized weight for concept (c) for technology (j)*
$w_{c,j,e}$	*importance rate for concept (c) for technology (j) given by expert (e)*
$w_{i,j}$	*normalized weight for metric (i) for technology (j)*
$w_{i,j,e}$	*importance rate for metric (i) for technology (j) given by expert (e)*

1. Introduction

According to the International Energy Agency (IEA, 2023a), energy demand is expected to rise persistently, emphasizing the pressing need to enhance energy security. There is also a need to accelerate the shift toward affordable, sustainable, and clean energy solutions (European Commission, n.d.). Reliability and energy security have also been underscored as critical factors concerning recently experienced vulnerabilities associated with the energy crisis, such as price volatility, lack of supply, interdependencies, and geopolitics (IEA, 2023a).

Technology roadmaps are widely used by governments and serve as guidance for decision-making in the energy transition (McDowall, 2012). Renewable energy sources (RES) are proposed to be key elements of decarbonization efforts (IEA, 2023b). However, RES intermittency poses a significant barrier to their widespread integration into the energy mix. In addressing this challenge, hydrogen (H_2) has emerged as a promising energy carrier that permits the use of surplus renewable electricity generation, such as that from photovoltaic and wind farms, thereby enhancing the flexibility of energy systems (Azzaro-Pantel, 2018). Currently, hydrogen is primarily produced and used in industrial applications, with on-site generation predominantly relying on Steam Methane Reforming (SMR) (IEA, 2023b). Substituting SMR production with electrolysis powered by renewable electricity produces renewable or "green" hydrogen. However, this transition entails trade-offs, such as weighing the reduction in CO_2 emissions against factors such as energy consumption, water usage, cost (González Palencia et al., 2022), safety risk, social cost-benefit, etc. (De-León Almaraz et al., 2022, 2014). Technical reliability might also be affected.

A concrete example is found in the European Union (EU) region, which has set the target of deploying a minimum of 40 gigawatts (GW) of renewable hydrogen through electrolysis by 2030 (European Commission, 2022). The EU then plans to produce up to 10 million tons of renewable hydrogen within its territory and import an extra 10 million tons from non-EU nations (European Commission, 2022). The technological readiness level (TRL) scale is a useful tool for evaluating the progress of a technology from the conceptual stage to market readiness (IEA, 2020). However, new hydrogen technologies are associated with uncertainty regarding their technological capabilities, including reliability. A disruption in energy supply chains could have a tremendous impact on society and the economy because it would affect all human activity (Franki and Višković, 2015). It is therefore crucial to identify the potential challenges and threats posed by a vulnerable HSC. This paper seeks to identify these issues.

The reliability of energy infrastructure is defined by McCarthy et al. (2007) and Scholten (2013) as "the *ability of [an energy] system to deliver the product (or service) transported over the network without interruption and without deterioration of its quality, i.e., the ability to supply the quantity and quality of energy desired by the customer when it is needed.*" In established systems, Reliability analyses (RAs) can be conducted at both component and system levels and are not constrained to evaluating technical aspects. Sustaining reliability encompasses a dual focus on technical and organizational considerations (Scholten, 2013).

Technically, reliability pertains to the probability that a system will accomplish its designated function without failure for a specified duration while operating under standard conditions (Blank, 2004; Ebeling, 1997). From an organizational perspective, the emphasis on maintaining reliability is directed at the various infrastructure companies or entities collectively responsible for asset management and operational oversight (Scholten, 2013). Depending on the scenario, reliability assessment can be undertaken using qualitative and quantitative methodologies.

As discussed in Section 2, while much research is taking place concerning the safety and reliability of specific hydrogen technologies or sub-systems, the technical reliability of hydrogen supply chains (HSCs) constitutes an important research gap. An HSC comprises several echelons (Fig. 1), including energy source, hydrogen production, transportation, storage, and distribution (e.g., HRS). Due to the diversity of options in the HSC, a single hydrogen supply chain will not exist.

The reliability of HSCs has been identified by De-León Almaraz et al. (2024) as a key issue that can affect the operation and growth of hydrogen technologies, social perception, and trust. Expectations can drive technological development by influencing stakeholders' perceptions of its future (Bakker et al., 2011; van der Duin et al., 2024). According to Bakker et al. (2011), making hydrogen visions credible requires analyzing HSC in its entirety, not only as separate technologies. However, analyzing reliability at the supply chain level is not a trivial task. The lack of necessary infrastructure and the interconnection of numerous new systems further complicate the ability to estimate the reliability of these networks (Kurtz et al., 2019). Moreover, in emerging supply chains, many aspects related to reliability can be challenging to assess within the actual infrastructure, creating challenges for developing industries (Turkcu and Tura, 2023). The development and implementation of emerging technologies require the involvement of a wide range of stakeholders and organizations (Ohlendorf et al., 2023). In addition, assessing the reliability of HSCs is a complex endeavor due to the prospective nature and associated uncertainty regarding technological advancement, demand projections (Park et al., 2022), locations, production capacities, inventory levels, and more. These arguments clearly illustrate the complexity of estimating the reliability of future HSCs.

Fortes et al. (2015) concluded that integrating socioeconomic stories with energy modeling enhances the reliability of energy scenario generation. Technological expectations are real-time representations of future technological situations and capabilities. That is, they involve a combination of appraisals of the expected progress of the technology, its future markets, and its societal context (Bakker et al., 2011; Borup et al., 2006). HSC stakeholders and society must deal with collective expectations in one way or another (Bakker et al., 2012).

The widespread utilization of hydrogen will require extensive research into the operation of a comprehensive, interconnected production, storage, and delivery network (Moradi and Groth, 2019). With this background, this paper aims to explore prospective methods for evaluating the reliability of new hydrogen supply chains in the year 2030. In this work, the problem definition consists of two options for green hydrogen production. As displayed in Fig. 1, the primary energy sources are solar and wind. The electricity generated from these sources will provide power for the electrolysis process. The hydrogen thus produced in gaseous form at low pressure (30 bar) can be stored at low, medium, or high pressure in pressurized containers or retained in liquid form (some applications may require the conversion of hydrogen into other chemical compounds, e.g., ammonia). Hydrogen delivery systems can be established to transport it to the point of use. The two research questions that drive the study are:

- RQ1: How can the reliability of the prospective hydrogen supply chain be analyzed?
- RQ2: What level of reliability can be anticipated for hydrogen supply chains by 2030?

The specific objectives are as follows:

- Identify or propose an appropriate method for assessing strategic-level HSC reliability.
- Gather data about systems and identify the expectations of experts and stakeholders.
- Calculate an HSC reliability index.
- Disseminate insights obtained from this methodology across various domains.

This study contributes to the body of knowledge on the reliability and future of low-carbon or green hydrogen supply chains by proposing mixed methods to facilitate the prospective evaluation of reliable hydrogen supply chains in 2030, as assessed by twenty-nine hydrogen experts. It helps bridge the gap between energy engineering research and social science with the complex integration of the five critical nodes in HSCs (energy source, production, transportation, storage, and refueling stations). Based on the literature review included in Section 2, the conceptual model proposed by McCarthy et al. (2007) serves as a

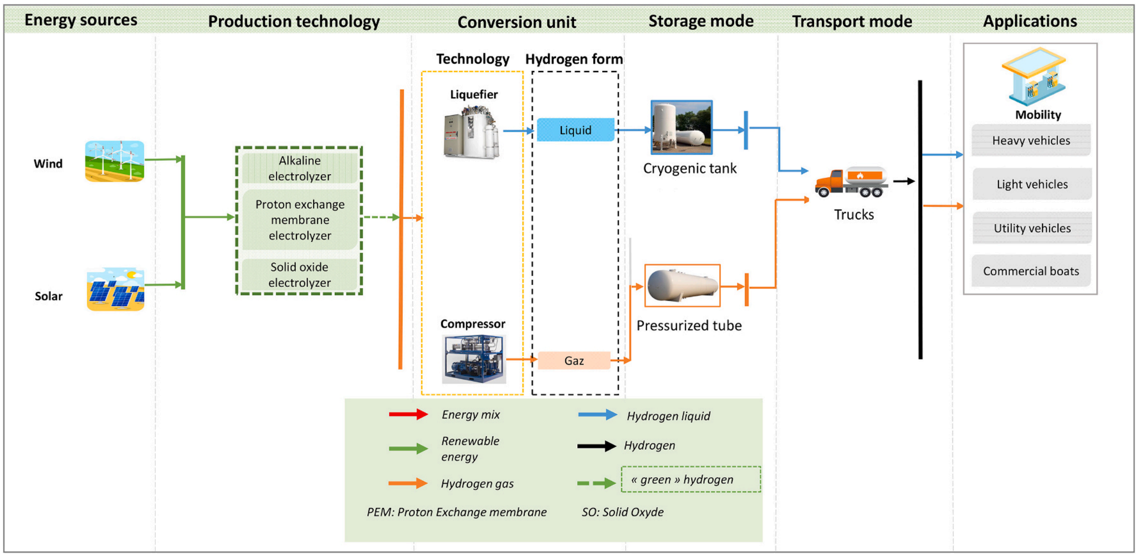


Fig. 1. Superstructure of hydrogen supply chains (Moustapha Mai et al., 2023).

reference for this work. In the quantitative part, multi-attribute utility theory (MAUT) (Leimeister and Kolios, 2018) is used to calculate the reliability index of the HSC. On the qualitative side, a novelty is the use of horizon scanning as an anticipatory method for capturing technological expectations (Jahel et al., 2023) for the next seven years (horizon: year 2030). Another methodological novelty is the data collected (using interviews and surveys) from multiple international stakeholders that captures several contextual perspectives (Sovacool et al., 2018). Although prospective studies of hydrogen technologies that consider expectations are not new (Budde et al., 2012; Köhler et al., 2010; McDowall, 2012; van Kerkhof et al., 2009; van Bree et al., 2010), the originality of this research lies in the fact that, according to the literature review conducted, there have been no studies dedicated to exploring expectations regarding technical reliability in the broader sense. An interdisciplinary team developed the current research to apply the various methodologies. This may have practical implications because using techniques to evaluate the potentiality of new technologies and systems is crucial for shaping technological progress toward sustainability (Hara et al., 2024).

This research is structured as follows: Section 2 provides a comprehensive literature review of the reliability assessment of hydrogen technologies. Section 3 introduces the methodology, detailing the specific methods, tools, reliability assessment, data collection approaches, and case study definition. Section 4 presents the results and discusses all the cases, identifying the study's strengths and limitations. Conclusions, contributions, and limitations are offered in Section 5. Finally, Section 6 discusses potential future directions.

2. Literature review: hydrogen and reliability

Both risk and reliability assessments are crucial prerequisites for building public trust in hydrogen infrastructure, especially given the challenges associated with addressing major hydrogen-related accidents (Scholten, 2013). The connection between risk and reliability assessments is strengthened by the fact that many of the tools employed in risk assessment are also utilized for evaluating reliability (De-León Almaraz et al., 2024). Due to hydrogen's inherent characteristics, safety risks are of a significant concern throughout the stages of production, transportation, storage, and utilization (Guo et al., 2021). In the near future, more risk assessment studies (Sharma et al., 2023) are expected to address the risks related to hydrogen facility construction, while "business interruption covers could gain importance as hydrogen becomes more embedded in the global economy, and the need to guarantee uninterrupted supply of renewable energy to fuel hydrogen production becomes ever-more pressing" (Swiss Re Institute, 2022). It is highlighted that financial and economic risks are also involved in deploying HSCs.

Safety risk and reliability assessments, while distinct, are interconnected (Clean Hydrogen Partnership, 2023; Tugnoli et al., 2009). Indeed, safety is a critical factor and is incorporated into international standards (e.g., ISO). However, safety risk can be seen as one of the many variables involved in gauging reliability (McCarthy et al., 2007). For instance, within the electricity sector, reliability encompasses the management of energy flows, network capacity, and equipment maintenance, as well as the mitigation of supply interruptions and congestion issues, as discussed by Scholten (2013). Section 2.1 presents methods used to assess reliability, and Section 2.2 discusses specific hydrogen reliability studies, including those that have calculated a reliability index. These outcomes also help identify gaps and potential methodological strategies.

2.1. Reliability assessment

There are many reliability methodologies, including quantitative, qualitative, and semi-quantitative ones. Leimeister and Kolios (2018) presented an exhaustive list of reliability assessment methods (Fig. 2). Quantitative reliability assessments include reliability evaluation of series and parallel systems, Markov analysis, Reliability Block Diagrams and Fault Tree Analysis, etc., based on the failure probabilities of any engineering system. Qualitative reliability assessments come into play when probabilities cannot be quantified due to data limitations. In such cases, a qualitative approach, often relying on expert opinions, can help to develop a reliability model (McCarthy et al., 2007; Seker and Aydin, 2022).

Recently, some efforts have been made to integrate the reliability dimension into analyzing energy systems and hydrogen subsystems. Hydrogen risk assessment models (HyRAM) can be used to assess the quantitative safety risk of hydrogen facilities based on past failure data (Ehrhart et al., 2021; Groth and Al-Douri, 2023; Hecht et al., 2021; West et al., 2022). Yue et al. (2021) critically reviewed various aspects of the hydrogen energy system, including its technologies, applications, trends, and challenges. Müller (2022) has documented the effect of reliability engineering on the performance of hydrogen technologies. Su et al. (2019) explored the operation of natural gas pipeline networks in terms of supply reliability and operation efficiency using multi-objective optimization. Kashanizadeh et al. (2022) used reliability indices to assess multi-energy systems that employ battery energy storage. Kurtz et al. (2020, 2019) have presented the failure rates and reliability growth for 29 hydrogen refueling stations (HRS) in California using failure modes and maintenance events (FMME) that were recorded in datasets of the National Fuel Cell Technology Evaluation Center.

Moreover, Wu et al. (2022) developed a reliability model for integrated electricity-gas systems, which uses hydrogen as a fuel, through Monte Carlo simulation. Condition-based monitoring effects on the reliability and resilience of multi-energy infrastructure systems are discussed in the work of Yodo et al. (2023). Chauhan et al. (2023) used human error assessment and a reduction technique (HEART) to assess the reliability of human performance (based on expert opinion) in HRSs. They proposed a framework that integrates the Bayesian network, best-worst method, and HEART techniques to enhance the safety and reliability of HRS's operation and maintenance practices. Additionally, Fetanat and Tayebi (2024) introduced a reliability-based strategy aimed at prioritizing hydrogen technologies for decarbonizing the oil refining industry. Their approach seeks to streamline decision-making processes in this critical area. In their study, the reliability assessment was conducted for the independent subsystems using CRITIC (combining criteria interaction through inter-criteria correlation) and CRADIS (compromise ranking of alternatives from distance to ideal solution) tools. To collect data, the authors developed a questionnaire, engaged in brainstorming sessions (Delphi method), and collected feedback from four experts from the oil industry.

2.2. Reliability index calculation for hydrogen systems

Rigorous investigation and quantification of the risk and reliability issues associated with hydrogen technologies are critical to ensuring both their wider adoption and safe, economical operation (Groth et al., 2024). Energy source technologies can stand out from the rest of the nodes in the HSC. Al-Douri and Groth (2024) developed and disseminated detailed description of electrolysis system and design documentation. Hydrogen storage and transportation play a vital role in the hydrogen supply chain. The basic method for obtaining liquid hydrogen involves a process of liquefaction, which necessitates the conversion of energy and the creation of harsh conditions. The safety problems associated with hydrogen arise from its flammability, low ignition energy, and its tendency to cause the embrittlement of metals (Wang et al., 2024). Hydrogen refueling stations (HRSs), including compressors, storage tanks, dispensers, and priority control panels, are crucial elements in maintaining safety. Leakage can lead to fires or explosions due to ignition sources (Kwon et al., 2022). The reliability can be defined for a HRS as the probability of a system performing its function without failure for a specified duration under standard conditions (Blank, 2004; Ebeling, 1997; Kurtz et al., 2019). The functionality failure of such

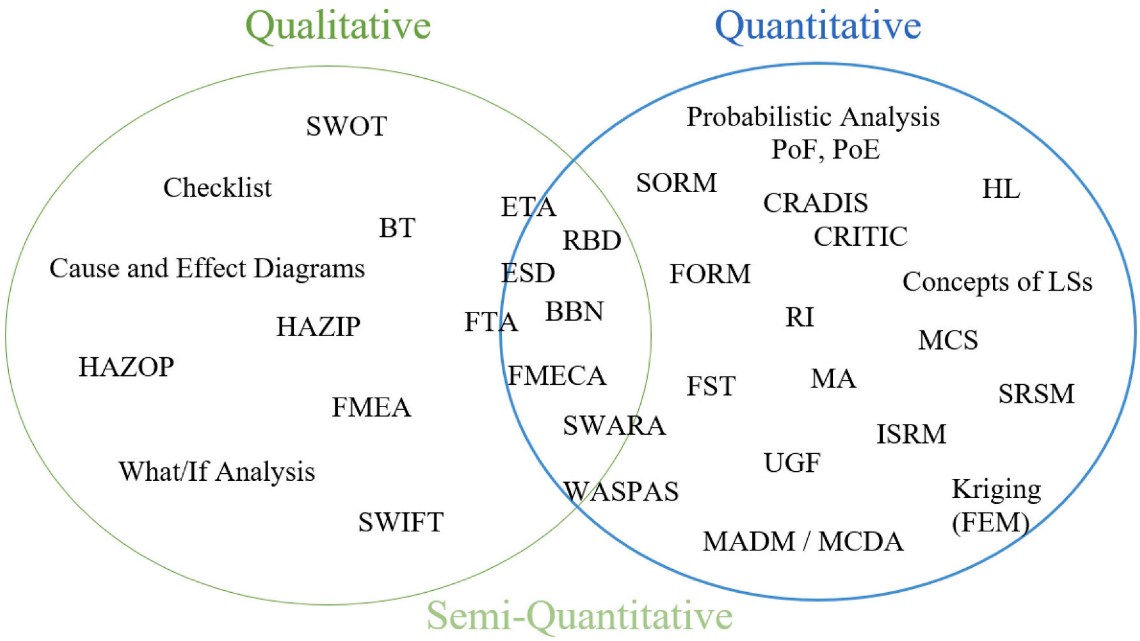


Fig. 2. Classification of methodologies[a] used to assess reliability (authors' creation modified from Leimeister and Kolios (2018) including methodologies from Chi et al. (2024); Correa-Jullian and Groth (2022); Fetanat and Tayebi (2024); He et al. (2022); Seker and Aydin (2022)).

[a]BBN: Bayesian (Belief) Network, BT: Bow-Tie (Analysis/Diagram), CRADIS: Compromise Ranking of Alternatives from Distance to Ideal Solution, CRITIC: Combining Criteria Interaction Through Inter-criteria Correlation, ESD: Event Sequence Diagrams, ETA: Event Tree Analysis, FEM: Finite Element Model Kriging method, FMEA: Failure Mode and Effect Analysis, FMECA: Failure Mode Effect and Criticality Analysis, FORM: First Order Reliability Method, FST: Fuzzy Set Theory, FTA: Fault Tree Analysis, HAZID: Hazard Identification, HAZOP: Hazard and Operability Studies, HL: Hasofer and Lind, ISRM: Importance Sampling Reduction Method, MA: Markov Analysis, MADM: Multi Attribute Decision Making, MCDA: Multi Criteria Decision Analysis, MCS: Monte Carlo Simulation, PoF: Probability of Failure, PoE: Probability of Exceedance, RAs: Reliability analyses, RBD: Reliability Block Diagram, RI: Reliability Index, SORM: Second Order Reliability Method, SRSM: Stochastic Response Surface Method, SWARA: Stepwise Weight Assessment Ratio Analysis, SWIFT: Structure What-if Technique, SWOT: Strengths, Weaknesses, Opportunities, and Threats, UGF: Universal Generating Function, WASPAS: Weighted Aggregated Sum Product Assessment.

systems may lead to catastrophic damages. Research on hydrogen network or supply chain reliability was reported to be at an early stage ten years ago (Scholten, 2013). Today, despite efforts to gauge safety risks for specific technologies, there remains a relatively limited exploration of the reliability of HSCs. Assessing the reliability of HSCs proves challenging due to the diverse array of stakeholders, such as producers, designers, managers, technology developers, policymakers, users, etc. (De-León Almaraz et al., 2024). Fortes et al. (2015) concluded that many studies use qualitative and quantitative modeling methodologies separately, primarily focusing on techno-economic quantitative insights while neglecting some social aspects that might be critical in decision-making.

Different ways exist to evaluate the reliability of hydrogen systems. For instance, technology maturity is related to the reliability of the system and can be presented in the form of TRLs, i.e., technology readiness levels (IEA, 2022) – Table 1. Although the NASA originally considered a 9-level TRL, the scale has been extended to a 11-level TRL in the IEA report (2022) to include the mature commercial use. However, many variables can make reliability evaluation different depending on the scope of assessment. Four variables were identified from the literature review: 1) reliability methodologies, 2) time horizon, 3) reliability analysis level, and 4) elements considered in the reliability index calculation (see Sections 2.2.1–4). In Table 2, research that presents a calculation of a reliability index for hydrogen technologies is listed and classified to aid in discussing the current research's main gaps and contributions.

2.2.1. Reliability analysis levels

Hydrogen technologies are arranged along a nested hierarchy of components, technologies, systems, and supply chains (macro level). Some studies evaluate specific technologies, like storage for liquid H_2 (LH_2) (Correa-Jullian and Groth, 2022) or compressed H_2 (CH_2) (Li et al., 2023). At a system level, most papers examine decentralized Power-to-X systems using RES, electrolysis with storage in gaseous form (Park et al., 2024a), and, in some cases, stationary fuel cells (Thakkar and Paliwal, 2024; Wu et al., 2022; Wu and Wang, 2023) or refueling stations (Elshurafa et al., 2022). In most system approaches, storage technology is included, but the transportation node is excluded. McCarthy et al. (2007) explored various options including renewable and fossil fuels as energy sources for hydrogen supply and studied its reliability considering production, and transportation, excluding storage.

Adopting a multi-level perspective in the assessment of the reliability of hydrogen supply chains might be useful for analyzing the related dynamics (van Bree et al., 2010) and identifying areas of concern. According to the literature review conducted, there is a gap related to

Table 1

Technology Readiness Levels – TRLs (IEA, 2020; IEA, 2022)*. TRL 0–4: concept or small prototype. TRL 5–7: Large prototype. TRL 7–8: demonstration. TRL ≥ 9: early adoption. TRL 11: mature. CH_2: Compressed hydrogen storage system, LH_2: Liquid hydrogen storage system, SMR: Steam Methane Reforming, PV: Photovoltaic energy, Wind: Wind energy.

Node	RES		Production		Storage		Transportation		Use
Technology	PV	Wind	Electrolysis	SMR	CH_2*	LH_2*	CH_2*	LH_2*	HRS*
TRL	≥9	≥9	≥9	11	11	11	11	7	9

Table 2

Reliability index calculation and its variables. New abbreviations: EDNS: Expected Demand Not Supplied, HS: Horizon Scanning, LOHC: Liquid Organic Hydrogen Carriers, LOHP: Loss of Hydrogen Probability, LOLP: Loss of Load Probability, LPSP: Loss of Power Supply Probability, NG: Natural Gas, PEM: Proton Exchange membrane, PtG: Power to Gas.

Reference	Reliability methodologies			Time horizon		Reliability analysis level		Technologies included				Comments
	Quantitative	Semi-qualitative	Techniques, methods or KPIs*	Present	Futuristic	Technology or component	System or HSC	Energy sources	Production	Transportation	Storage	
McCarthy et al., 2007	X	X	MADM	X		Some nodes	HSC	PV, Wind, Imported NG	Electrolysis, SMR	CH2 - Pipeline		
He et al., 2022	X		PoE, LPSP*	X			System	PV, Wind	Electrolysis		CH2	Includes battery, thermal energy, and pumped hydro storage system
Liu et al., 2022	X		Probabilistic: LPSP*	X			System	PV, Wind	Electrolysis		CH2	
Wu et al., 2022	X		Probabilistic: LOLP*, LPSP*, EDNS*	X			System	RES	Electrolysis, methanation		CH2	Application to power-to-hydrogen-heat-methane system
Seker and Aydin, 2022	X	X	MCDA, FST, SWARA, WASPAS	X		Production		H2S**	Thermochemical, Electrochemical, Photochemical, and Plasma			Reliability, sustainability, efficiency, operational suitability, and technical maturity criteria have been considered
Correa-Jullian and Groth, 2022		X	FMEA, ESD, FTA	X		Storage					LH2	Qualitative analysis of liquid hydrogen.
Elshurafa et al., 2022	X		EDNS		X		System	PV, Wind	Electrolysis		CH2	Application to decentralized system
Wu and Wang, 2023	X		MCS	X			System	RES	Electrolysis		CH2	Application to PtG
Li et al., 2023	X		FEM	X		Storage		–			CH2	
Park et al., 2024b	X		LOHP*	X			System	PV, Wind	Electrolysis (PEM)		CH2	
Fetanat and Tayebi, 2024	X		MADM, CRITIC, CRADIS	X		Production		PV, Wind, NG, biogas, nuclear	Electrolysis, SMR			Industrial applications. Each technology is evaluated and compared.
Thakkar and Paliwal, 2024	X		MADM, LOLP*, LPSP*	X			System	PV, Wind	Electrolysis		CH2	Application to hydrogen storage system with fuel cell
Chi et al., 2024	X		UGF		X		System	PV, Wind	Electrolysis			Application to power sector
Park et al., 2024a	X		LOHP*		X		System	PV, Wind	Electrolysis (PEM)		CH2	Industrial application considering fuel cell.
Current study	X	X	MADM, HS, Interviews, Survey	X	X	5 nodes (including use)	HSC	PV, Wind, NG	Electrolysis, SMR	CH2 - Tube trailers, LOHC	CH2, LOHC	Application: Industry, HRS. LH2 not included due to the lack of experts

integrating all the hydrogen supply chain nodes. Transportation is usually almost disregarded, and storage is not included in McCarthy et al.'s (2007) work. In contrast, all the nodes of the hydrogen supply chain are considered in the current work.

2.2.2. Reliability methodologies

According to Leimeister and Kolios (2018), categorization and quantitative methods are identified as the most popular reliability approaches, as shown in Table 2. Quantitative methods include probabilistic analysis, Monte Carlo simulation, fuzzy set theory, Markov analysis, multi-attribute decision-making, and multi-criteria decision analysis. Fazli-Khalaf et al. (2020) proposed a mathematical model for helping maximize HSC reliability by quantitatively identifying reliable and unreliable hydrogen production plants. The model implements primary and backup plans to ensure maximum demand satisfaction in both normal and disruptive scenarios. The study of Correa-Jullian and Groth (2022) used classic reliability qualitative methods (FMEA, ESD, and FTA).

In addition, McCarthy et al. (2007) and Seker and Aydin (2022) used semi-qualitative approaches and, more specifically, mixed methods. In particular, McCarthy et al. (2007) used MADM and focus groups, and Seker and Aydin (2022) employed MCDA, fuzzy set theory, and interviews. Several pieces of work (McCarthy et al., 2007; Thakkar and Paliwal, 2024) have employed multi-attribute or multi-criteria decision-making methods, incorporating multiple criteria or aspects in their analysis. However, in some cases, studies do not solely measure reliability criteria (Fetanat and Tayebi, 2024; Seker and Aydin, 2022).

Quantitative methods provide a more accurate approximation of the reliability and risk of specific hydrogen technologies and are an important tool for enabling the safe deployment of many engineering systems. However, quantitative studies require reliable data, which are currently lacking concerning the expanding applications of hydrogen systems (Groth et al., 2024). Moreover, the perception and expectation of the reliability of hydrogen technologies and their supply chains might affect decision-making concerning hydrogen technologies. Using mixed methods and specific qualitative techniques for data collection can address this situation.

2.2.3. Time horizon

Since many studies are quantitative, they use current data for evaluations. In some cases, scenario analysis is performed with hourly, monthly, or annual time horizons. These works are categorized in Table 2 as having a present time horizon. Two studies consider futuristic long-term scenarios while considering specific scenarios using quantitative methods (Chi et al., 2024; Park et al., 2024b). While multi-period scenario analysis is of utmost importance, it is also critical to know the expectations of hydrogen experts because R&D or investment activities may be based on these. Roadmaps have been presented for the next two decades (European Commission, 2020), making it necessary to adopt a medium- to long-term perspective on technology reliability. In the next few years, the perceived reliability of hydrogen technologies is liable to affect both scale-up efforts and commercialization. Technological readiness levels differ between technologies (IEA, 2022), and many research projects that are being supported worldwide, making it uncertain which options will be considered most reliable in the future.

Quantitative methods create a more accurate picture of the risk and reliability of hydrogen technologies, but how people think about the latter and what they expect requires more research. When actors talk about innovations or transition more broadly, they convey meaning, shape categories, and (co-)create expectations (Bakker et al., 2011; Ohlendorf et al., 2023). Qualitative anticipatory methods refer to the future without considering numerical indicators (Jahel et al., 2023). They usually provide representations of the future in the form of narratives and are implemented in a participatory way, making it possible to integrate different disciplinary perspectives (Jahel et al., 2023). This study uses qualitative and quantitative methods to holistically assess the reliability of prospective hydrogen supply chains in 2030.

2.2.4. Elements of the reliability index calculation

The aspects included in reliability assessment vary from one study to another. Liang and Pirouzi (2024) connect reliability and flexibility. Fetanat and Tayebi (2024) include technological operation and maintenance measures, such as inspection, repair, lifetimes, service life, and fault tolerance. The reliability index is often calculated based on the likelihood of hydrogen loss (LOHP) (Park et al., 2024a,b), the likelihood of power supply loss (LPSP) (He et al., 2022; Liu et al., 2022; Wu et al., 2022), and, in more complicated systems, the likelihood of expected demand not being met by the power systems (EDNS), the likelihood of lost renewable energy (LORP), the likelihood of expected renewable energy not being supplied by renewable generators (ERNS), and the likelihood of gas loss (LOGP) and expected gas not being supplied (EGNS) (Wu et al., 2022). McCarthy's model (2007) considers reliability using 20 metrics. The authors take a macro-level approach that uses mixed methods and considers technical and organizational aspects of reliability.

Although the work of McCarthy et al. (2007) was developed almost twenty years ago, it considers the largest set of elements in evaluating hydrogen systems. However, the proposed metrics may require validation.

From the literature review, four gaps have been identified (Fig. 3). First, there is a need for a comprehensive reliability evaluation of HSCs that includes all critical nodes. Second, effective data collection methods are required for assessing prospective HSCs. Third, employing mixed methods is essential for studying the reliability of these systems. Finally, there is a need to identify key categories for calculating a reliability index.

2.3. Proposed methodology for assessing the reliability of future HSCs

Based on the conducted literature review, the research of McCarthy et al. (2007) was the first to explore hydrogen reliability for different pathways, encompassing multiple nodes within the supply chain using multi-attribute utility theory (MAUT). This work incorporated a methodology with a comprehensive list of variables, aligning with conventional reliability studies but tailored for application to larger and more complex systems. Although the methodology relies on subjective estimations provided by experts rather than exact measurements of these variables, the work can be considered an initial attempt to integrate both technical and organizational aspects of reliability. The reliability assessment was developed for the electric power generation system, and reliability was defined by including two elements: adequacy and security (McCarthy et al., 2007). Adequacy refers to the system's ability to meet customer needs under normal working conditions. In contrast, security involves the system's response to unanticipated disturbances (McCarthy et al., 2007).

Together, adequacy and security delineate the overall reliability of the system, broadly described as its ability to supply the quantity and quality of energy desired by the customer when needed (McCarthy et al., 2007). The authors used focus groups to collect data and rate two hydrogen pathways. At first glance, McCarthy's approach may be extended to evaluate supply chain reliability. Elaborate assessments that determine the reliability of the supply chains can reduce uncertainty, instill trust among stakeholders, and provide a more robust foundation for action plans, resource allocation investment, and policy implementation. However, the RQs from this work cannot be completely addressed using the methodology of McCarthy et al. (2007) because a prospective and international assessment was not developed in their study.

Specific methodologies and tools may be required to facilitate discussion and collaboration among the many hydrogen stakeholders (van Kerkhof et al., 2009). In futures studies (Sacio-Szymańska et al., 2016) participatory foresight tools are a critical driver of economic innovation,

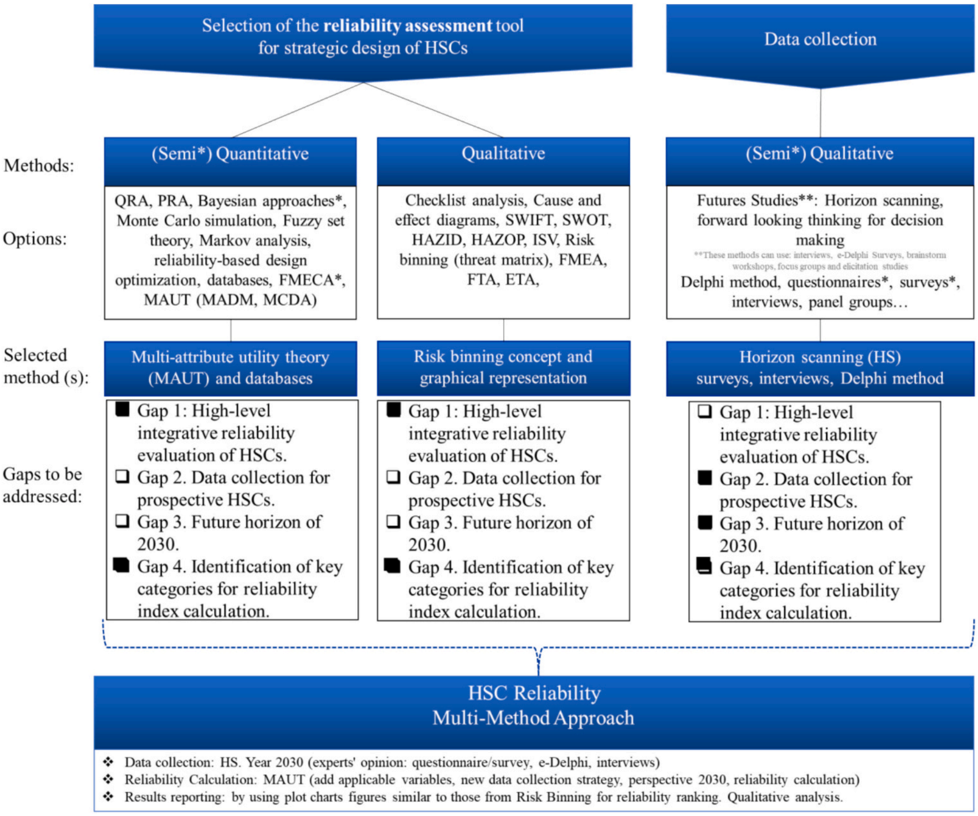


Fig. 3. Methodological underpinning (authors' creation).

offering practical benefits for corporate entities and policy decision-makers. This approach combines evidence-based quantitative methods with tailored qualitative approaches aligned with societal needs. Horizon Scanning (HS) is a goal-oriented process that integrates multiple methodologies to explore various future perspectives (van der Duin et al., 2024). Rather than making predictions, HS systematically examines future trends, encompassing both positive and undesired possibilities (Gáspár et al., 2021; Hideg et al., 2021). The methods typically used in this process include literature monitoring and processing, Delphi methods, brainstorming, data analysis, trend calculations, modeling, and simulations using info-communication technology and internet data sources (Raford, 2015).

It should be emphasized that for the macro-level prospective assessment of HSC's reliability, technological and organizational reliability should be considered. While there is no precise understanding of how to implement a realistic prospective reliability assessment, the involvement of individuals with valuable experience can enhance the behavioral realism of energy models (Krumm et al., 2022), and technological expectations can impact the deployment of predefined systems (Bakker et al., 2011; Borup et al., 2006). This typically involves combining methods, such as translating insights from empirical approaches into the model under consideration. Mixed methods refer to the integration of quantitative and qualitative research methods within a single study (Sovacool et al., 2018). The human element plays a pivotal role in economic science, particularly in the realm of investment decisions. Additionally, emotional factors may influence energy-related decisions and behaviors (Brosch et al., 2014). Energy-relevant choices and behaviors do not take place in isolation but in the context of markets and political systems, which interact with the more proximal determinants of decision-making (Brosch et al., 2016). In this work, stakeholders' opinions and expectations are considered of utmost importance for identifying their understanding of the reliability of the different nodes of the HSC. As an illustrative example, Kurtz et al. (2020) reported that data associated with an HSC operation and maintenance show that reliability and throughput are significant contributors to the price of hydrogen per kilogram information crucial for decision-making.

In summary, the possibility of using mixed methods (MAUT+HS) to assess the reliability of potential futures for green hydrogen supply chains in 2030 has been identified, with Fig. 3 and Table 2 highlighting the novel elements in this study. The proposed methodology, with the potential application of each tool and the indication of the gaps covered, is displayed in Fig. 3. In this work, reliability expectations are presented using graphical representations developing a risk binning matrix.

3. Methodology

This research relies on mixed methods (Fig. 3) by following three steps. The first step consists of the development of a quantitative framework that uses an additive approach for multi-attribute utility models, contingent upon the level of independence exhibited among the attributes adopted (Section 3.1). Multi-attribute utility theory (MAUT) is part of multi-attribute decision-making (MADM). It is also related to multi-criteria decision analysis (MCDA), which is classified as a sophisticated quantitative reliability-based method (Leimeister and Kolios, 2018). This methodology can support selecting the best option in the presence of multiple criteria, so it applies to the defined problem and to achieving one of the main objectives of this research, i.e., assessing strategic-level HSC reliability. The second step uses (semi) qualitative methods for the data collection process, particularly interviews, and surveys, which are implemented using the perspective of horizon scanning (ref. the year 2030) (see Section 3.2). The third step consists of the qualitative and quantitative analysis of the results (Section 4). Steps 1 and 2 are not sequential but developed in parallel.

When using HS, it is essential to remain cautious about potential expert biases and the distortion of assumptions caused by technology trends. As a preparation for methodological implementation, strategies are proposed to manage potential cognitive biases.[1] As noted by Bonaccorsi et al. (2020), research institutes and manufacturers hold high expectations for hydrogen vehicles and fuel cells, yet their technological promises remain unfulfilled. This study aims to counteract overoptimistic biases associated with the expected reliability of hydrogen technologies.

The conceptual framework McCarthy and colleagues used to assess reliability has been adapted, and new metrics have been proposed (i.e., size, territorial conditions, decentralization degree, and resources). For the HSC presented in Fig. 1, each node is first assessed separately, considering interdependencies with other technologies, and then the inputs are integrated into a mathematical model. Interviews and surveys are conducted individually (one expert at a time), not in focus groups, as in McCarthy et al. (2007). This generates deeper answers and justifications and reduces the biases that can arise in group panels (conformity, ingroup, outgroup, and false consensus biases) (Bonaccorsi et al., 2020; Korteling et al., 2023). The use of mixed methods mitigates framing and overconfidence biases (Bonaccorsi et al., 2020; Korteling et al., 2023), and the Delphi method allows the experts to rethink their original answers. The ambiguity effect (Korteling et al., 2023) is treated by preparing a Likert-scale grid with examples for different levels per technology.

The "H_2 reliability project" was developed through collaborative efforts across multiple disciplines, involving three research groups and five team members with complementary backgrounds: a chemical engineer from the Chemical Engineering Laboratory at the University of Toulouse, an energy systems engineer from the University of Corsica Pascal Paoli; a reliability expert (mechanical engineer) from the Suresh Gyan Vihar University, and two industrial engineers from the Institute of Operations and Decision Sciences of Corvinus University of Budapest (one specializing in risk assessment and the other in hydrogen supply chains). This multidisciplinary approach allowed the project to leverage complementary expertise, essential for addressing the complex challenges inherent in hydrogen reliability assessment.

Experts' diversity is crucial in managing various biases, including anchoring, optimism, pessimism, desirability, wishful thinking, overconfidence, and advocacy (Korteling et al., 2023; Bonaccorsi et al., 2020). To address affective forecasting (Bonaccorsi et al., 2020), a minimum number of experts per technology is established. This research, conducted in 2023, set a close time horizon for evaluation (2030) to mitigate the effects of foresight bias and the planning fallacy (Bavaresco et al., 2020; Korteling et al., 2023).

3.1. MAUT for the reliability evaluation of hydrogen supply chains

This section deals with the application of MAUT for the reliability evaluation of HSC. The reliability of HSC is defined in terms of two components, which are known as General Objectives (GOs): Adequacy and Security (McCarthy et al., 2007). HSC reliability can be broadly described as the ability to supply the quantity and quality of energy desired by the customer when needed. The HSC is built on a structured hierarchy (Fig. 4), consisting of various subsystems or nodes (j) representing technologies related to energy sources, production, storage, transportation, storage, and market.

3.1.1. Attributes, concepts, and metrics related to reliability assessment

Based on the models developed in (McCarthy et al., 2007), the researchers involved in this project revised and selected the most relevant attributes, objectives, concepts, and metrics. A list of two general objectives (Adequacy and Security), five concepts (Capacity, Flexibility, Infrastructure, Consequences, and Energy Security), and 20 metrics proposed by McCarthy et al. (2007) plus four new metrics are involved (Table 3 and Fig. 4). One noteworthy observation is the varying impact of the metrics used in the assessment for each technology. The new metrics introduced in the proposed model are system size, degree of decentralization, territorial constraints, and resources and operations. At this stage, the first variable was incorporated into the capacity concept, while the other three are considered as the part of flexibility. It is acknowledged that this categorization may need further discussion.

The relationship between the attributes is displayed in Fig. 4. First, the five subsystems (j) are connected to the two *General Objectives (Adequacy and Security)* to indicate that each subsystem will undergo an independent reliability evaluation. The *GOs* (r) are divided into *Concepts* (c), and each concept is subsequently divided into various *metrics* (i). The division allows appropriate measures to be developed for the different components/subsystems of the HSC system. Metrics that are only applicable to the energy source nodes (solar and wind) are those of energy security. For this reason, the metric of *resources and operations* was included, allowing the analysis of this aspect for all the nodes.

The definition of each metric was crucial in defining the scales (Table 3). For each metric, the definition was developed first, and then the research team developed a description of different levels of reliability. First, a three-level Likert scale was proposed for reliability and tested by two hydrogen experts. The feedback was that low, medium, and high-reliability assessments did not apply to all metrics and technologies. The experts approved a second test that employed five reliability levels. Consequently, in the final survey, the Likert scale relies on a five-reliability level matrix (1–5 Likert scales, Table 4) with examples for each metric and each supply chain node to increase the validity of the survey.

3.1.1.1. Utility per metric (i).

This formulation defines comprehensive utility ($u_{i,j}$) by evaluating each reliability metric (i) for each technology subsystem (j). In this sense, utility is used to explain reliability. This was done by collecting the experts' opinions (e) using a five-item Likert survey. The target population for the survey was experts associated with the different nodes of the HSC. The Likert scale allows the assignment of quantitative values for each position using a qualitative rating scale to streamline statistical analyses while capturing attitudes. In this approach, values ranging from 1 to 5 are assigned to qualitative positions to reflect the attributes' reliability and importance. A rating of '1' indicates high reliability and low importance, whereas '5' signifies poor reliability and high importance (Table 4). Consequently, in the context of weighted and aggregated utility scores, higher scores indicate lower reliability, while lower scores represent greater reliability.

3.1.2. Calculation of the reliability index using MAUT

MAUT was used to integrate utility (reliability) and importance ratings, creating reliability indices that reflect each subsystem's concepts and overall objectives. In this work, the attributes are assumed to be independent to permit the utilization of the additive formulation of the model. The reliability indices are then aggregated to establish comprehensive indices for the HSC, similar to the hierarchical combination of subsystems. Prior to aggregation, the ratings given in the survey (originally using a 1–5 scale) are normalized using the additive MAUT approach. Using the tested scale, utility ratings are divided by five (highest Likert scale value), converting them into a decimal scale from 0 to 1. In parallel, importance ratings are divided by the sum of importance ratings for the specific set of subordinate attributes they pertain to. This transformation converts importance ratings into weights representing the proportion of a higher-level attribute described by its subordinates while preserving the necessary relationships. This transforms the importance ratings into weights related to the proportion of a higher-level attribute described by each of its subordinates and maintains the requisite that $\sum_{i=1}^{n} w_i = 1$ for a set of "n" attributes.

---

[1] Cognitive biases refer to systematic, universally occurring tendencies, inclinations, or dispositions in human decision-making that may make it vulnerable to inaccurate, suboptimal, or wrong outcomes (Korteling et al., 2023). Biases can affect how people think, and judge things.

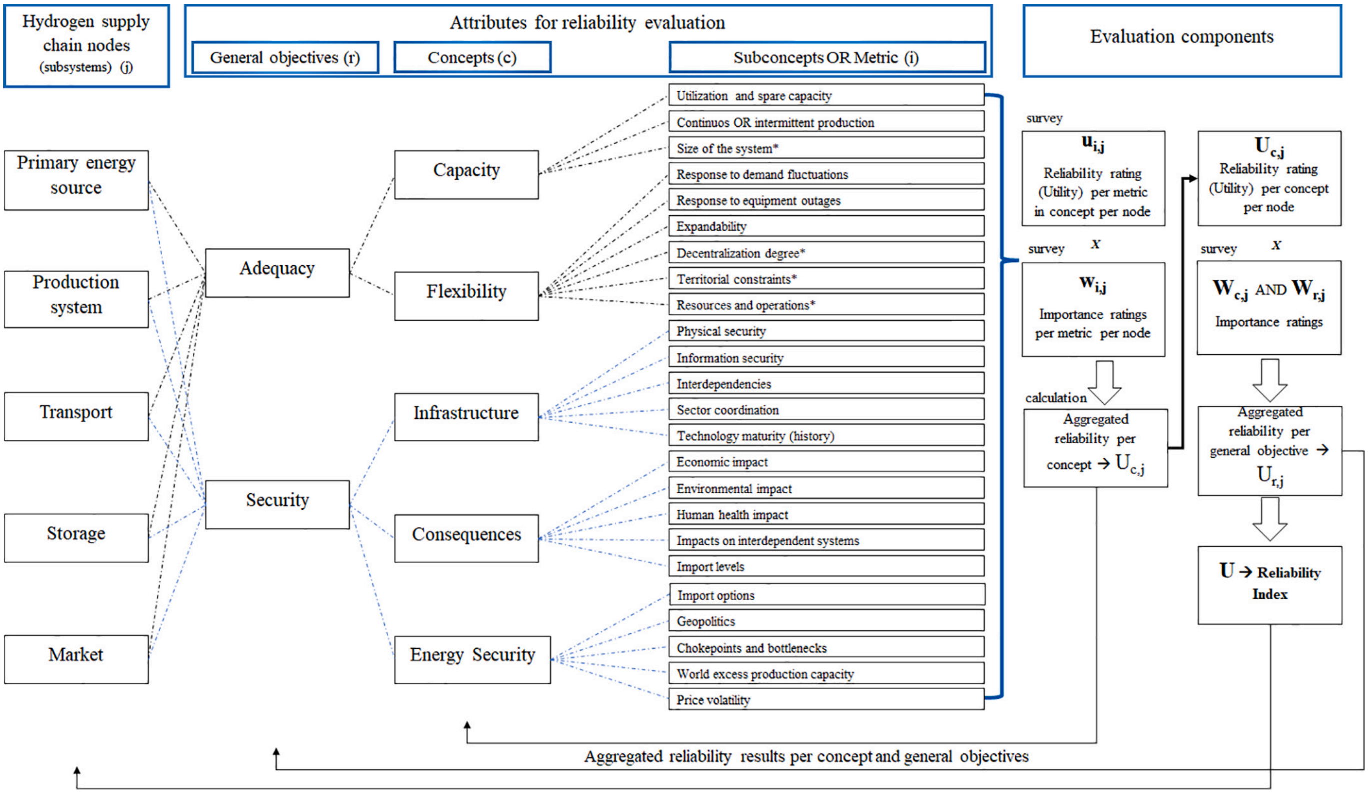


Fig. 4. Reliability assessment framework for hydrogen supply chains (authors' creation based on McCarthy et al. (2007)).

As proposed by McCarthy et al. (2007), expert ratings underwent three rounds of aggregation using MAUT to generate diverse reliability indices. Initially, within a specific subsystem, utility ratings for the metrics were merged with the metric importance weights, producing utility scores at the conceptual level. These scores were then aggregated with concept importance weights to establish utility scores for the broad objectives of adequacy and security applicable to every subsystem. Subsequently, utility scores related to a GO were integrated across the subsystems, culminating in an encompassing utility score for the entire HSC pathway, i.e., the reliability index. In this sense, there is a difference between the calculation of utilities at different levels of aggregation (i.e., $u_{i,j,e}$ (see Section 3.1.1)/$u_{i,j}$/$u_{c,j}$/$u_{r,j}$) as developed below.

Following McCarthy et al.'s (2007) methodology, the aggregated evaluations from all experts (N) for a specific technology are processed by averaging across each metric. Subsequently, the values are normalized to a scale from 0 to 1. The normalized $u_{i,j}$ is calculated as follows:

$$u_{i,j} = \frac{\frac{\sum_e u_{i,j,e}}{N_{e,j}}}{MaxLikertValue} \tag{1}$$

where:

$u_{i,j,e}$: Reliability rate for metric (i) for technology (j) given by expert (e)

$N_{e,j}$: Total number of experts (e) evaluating technology (j)

MaxLikertValue: Highest allowed value in the survey: 5

$u_{i,j}$: Normalized utility for metric (i) for technology (j)

3.1.2.1. *Utility per concept (c).*

The utility functions are determined by the utilities associated with individual attributes and their respective weights. To calculate the utility per concept ($u_{c,j}$), two elements are needed: (1) $u_{i,j}$, and (2) the importance rates for metric (i) for technology (j) ($w_{i,j}$). The data collection process is explained in Section 3.2. The importance weights reflect the degree to which attributes lower in the hierarchy contribute to the attribute above them, relative to each other:

$$u_{c,j} = u_{i,j} w_{i,j} \tag{2}$$

$$\overline{w_{i,j}} = \frac{\sum_e w_{i,j,e}}{N_{e,j}} \tag{3}$$

$$w_{i,j} = \frac{\overline{w_{i,j}}}{\sum_i \overline{w_{i,j^*}}} \tag{4}$$

$$\sum_i w_{i,j} = 1 \tag{5}$$

where:

$u_{c,j}$: Utility per concept.

$w_{i,j,e}$: Importance rate for metric (i) for technology (j) given by expert (e)

$N_{e,j}$: Total number of experts (e) evaluating technology (j)

$\overline{w_{i,j}}$: Average importance rate for metric (i) for technology (j)

$\overline{w_{i,j^*}}$: Average importance rate for metric (i) included in concept (c) for technology (j)

$w_{i,j}$: Normalized weight for metric (i) for technology (j). The sum of weights should be 1 as displayed in Eq. (5).

3.1.2.2. *Utility per general objective (r).*

The final step is aggregating the utility for the general objectives (r). It is important to mention that adequacy and security are kept independent so that their impact can be analyzed separately. The related equations for this calculation are:

$$U_{r,j} = u_{c,j} w_{c,j} \tag{6}$$

$$\overline{w_{c,j}} = \frac{\sum_e w_{c,j,e}}{N_{e,j}} \tag{7}$$

Table 3
Definition of reliability variables included in the study[a] (authors' creation based on McCarthy et al. (2007)).

Variable	Definition
Reliability	Together, adequacy and security describe the system's overall reliability, broadly described as the ability to supply the quantity and quality of energy desired by the customer when needed.
General objectives	
Adequacy	This refers to the system's ability to supply customer requirements under normal operating conditions. It considers the system statically.
Security	Includes the dynamic response of the system to unexpected interruptions and relates its ability to endure them.
Concepts	
Capacity	The ability of the SYSTEM to provide sufficient throughput (product volume) to supply final demand
Flexibility	Degree to which the SYSTEM can adapt to changing conditions
Infrastructure vulnerability	Degree to which the SYSTEM is susceptible to disruption
Consequences of infrastructure failure	Degree to which a disruption in the SYSTEM causes harm
Energy security	Uninterrupted availability of energy sources at an affordable price
Metrics	
Utilization and spare capacity	Degree to which the capacity of the SYSTEM is being used
Continuous OR intermittent production	Degree to which the productivity of the SYSTEM is constant
Size of the system*	Degree to which the throughput volume can be modified in the SYSTEM
Response to demand fluctuations	Degree to which the SYSTEM is able to adapt to varying demand levels and locations
Response to equipment outages	Degree to which the SYSTEM can continue reliably operating in the event of equipment downtime (failure)
Expandability	Degree to which the SYSTEM can easily and cost-effectively be expanded
Decentralization degree*	Potential of the SYSTEM to be located anywhere in a given territory and be disconnected from other hydrogen technologies in the supply chain
Territorial constraints*	Degree to which it is possible to install the SYSTEM anywhere from a geographical perspective (e.g., due to resource availability, regulations)
Resources and operations*	Degree to which the human and material resources are needed in the SYSTEM to allow its operation (e.g., automation level, scarce materials)
Physical security	Degree to which assets in the SYSTEM are secure against physical threats (menaces)
Information security	Degree to which information assets in the SYSTEM are secure against threats (menaces)
Interdependencies	Degree to which the SYSTEM relies on other infrastructures for its reliable operation and is vulnerable to their disruption
Sector coordination	Degree to which coordination between stakeholders within the sector of the SYSTEM results in an effective exchange of information alerting stakeholders to emerging threats (menaces) and mitigation strategies
Technology maturity (history)	Degree to which the SYSTEM has been susceptible to disruption (failure) in the past
Economic impact	Degree to which a disruption in the SYSTEM causes economic damage to industry stakeholders, the government, or the public
Environmental impact	Degree to which a disruption in the SYSTEM causes environmental damage
Human health impact	Degree to which a disruption in the SYSTEM harms the health of employees and/or the public
Impacts on interdependent systems	Degree to which a disruption in the SYSTEM causes damage to interdependent systems
Import levels**	Degree to which the SYSTEM relies on resources from outside of the country
Import options**	Degree to which imports of energy sources used for the SYSTEM are concentrated among a small group of supplying countries
Geopolitics**	Degree to which the political and social conditions threaten the operation of the SYSTEM in relation to the supply of energy products
Bottlenecks (chokepoints)**	Degree to which imported energy resources are vulnerable to being blocked in narrow shipping channels or busy transportation lines
World excess production capacity**	Degree to which excess production capacity of the energy resources used in the SYSTEM exists in the global market
Price volatility[b],**	Degree of fluctuation in the average price of primary energy used in the SYSTEM

[a] Definitions as presented in McCarthy (2004) and McCarthy et al. (2007) with the exception of the variables denoted with **. Definitions of new variables (*) were developed by the research team.

[b] Although McCarthy et al. (2007) included price volatility as one metric of reliability, they also concluded that this can be considered an outcome rather than a feature. **In this approach, the energy security metrics can only be applied to the energy sources node (solar and wind).

$$w_{c,j} = \frac{\overline{w_{c,j}}}{\sum_{c} \overline{w_{c,j^*}}} \tag{8}$$

$$\sum_{c} w_{c,j} = 1 \tag{9}$$

where:

$U_{r,j}$: Utility per general objective

$w_{c,j,e}$: Importance rate for concept (c) for technology (j) given by expert (e)

$N_{e,j}$: Total number of experts (e) evaluating technology (j)

$\overline{w_{c,j}}$: Average importance rate for concepts (c) for technology (j)

$\overline{w_{c,j^*}}$: Average importance rate for concepts (c) included in objective (r) for technology (j)

$w_{c,j}$: Normalized weight for concept (c) for technology (j). The sum of weights should be 1 as displayed in Eq. (9).

3.2. *Data collection strategy*

A part of the proposed methodology is qualitative (Fig. 5). Horizon scanning involves completing questionnaires, employing participatory methods, conducting future workshops, and more, based on the insights of stakeholders, experts, policymakers, and decision-makers (Hideg et al., 2021). Questionnaires are one of the qualitative methods most often used (Bavaresco et al., 2020; Chauhan et al., 2023; Fetanat and Tayebi, 2024; McCarthy et al., 2007). The strategy involved developing a survey that was distributed and applied with a deductive research interview to collect experts' feedback.

As presented in Fig. 3 and Section 3.1, the calculation of the reliability index required data collection directly using expert judgments and other applicable reports. This is because some HSC technologies have not achieved technical maturity (see Sections 1 and 2). Scenarios for technology improvement and an increase in HSC installed capacity are expected in the coming decades. Learning the experts' opinions about the projected reliability of the HSC components for 2030 seemed appropriate. By using HS, the potential reliability index for 2030 was assessed. This prospective approach is different from that followed by McCarthy et al. (2007), in which the “present” perspective about the hypothetical network of refueling stations in the Sacramento (California) area was used. Another difference from this research is the data collection method, involving developing twenty-nine individual interviews and surveys. In contrast, McCarthy et al. (2007) worked with a

Table 4
Scale used to rate reliability and importance metrics[a] (authors' creation).

	1	2	3	4	5
Reliability	Very high	High	Medium	Low (Moderately poor)	Very low (poor)
Importance	Very low	Low	Medium	High	Very high

[a] Zero values were applied when the expert did not know the answer or thought that the question was not applicable.

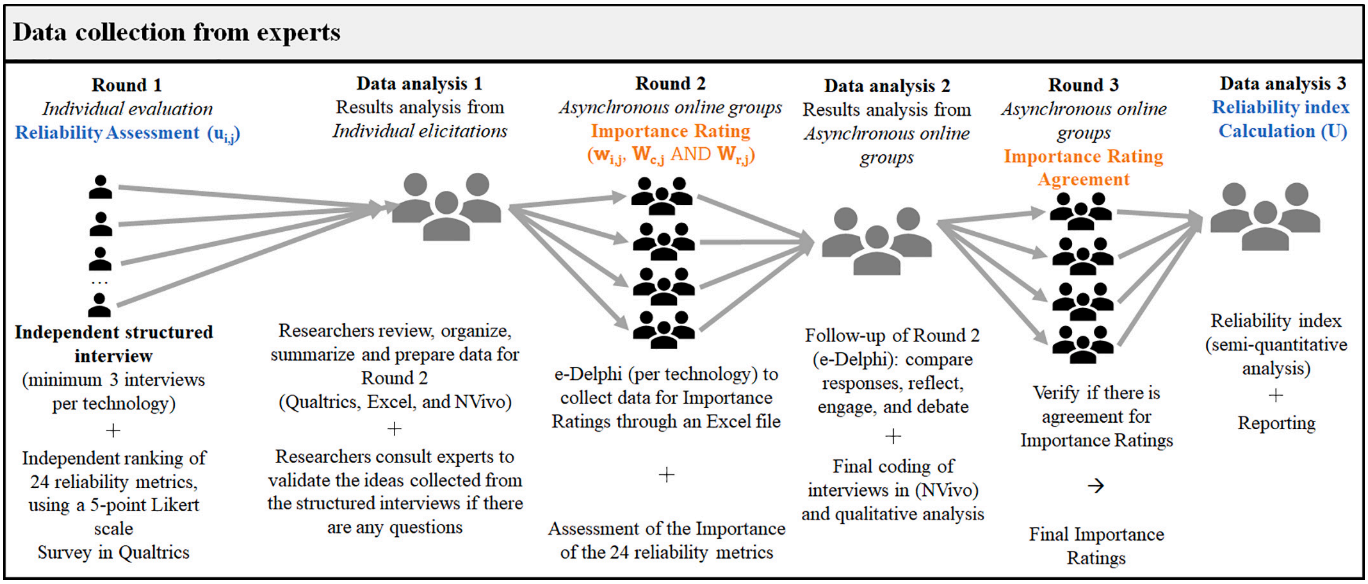
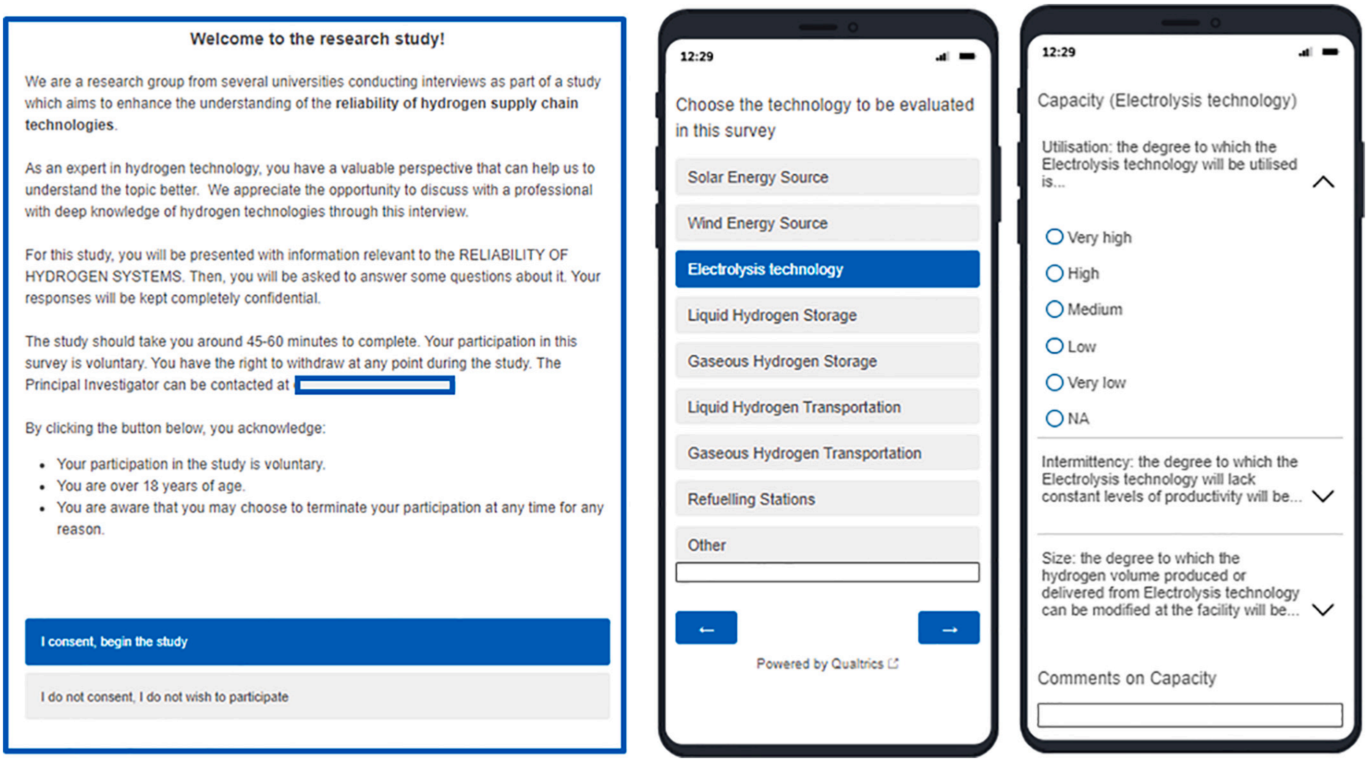


Fig. 5. Data collection methodology (authors' creation).



Welcome to the research study!

We are a research group from several universities conducting interviews as part of a study which aims to enhance the understanding of the **reliability of hydrogen supply chain technologies**.

As an expert in hydrogen technology, you have a valuable perspective that can help us to understand the topic better. We appreciate the opportunity to discuss with a professional with deep knowledge of hydrogen technologies through this interview.

For this study, you will be presented with information relevant to the RELIABILITY OF HYDROGEN SYSTEMS. Then, you will be asked to answer some questions about it. Your responses will be kept completely confidential.

The study should take you around 45-60 minutes to complete. Your participation in this survey is voluntary. You have the right to withdraw at any point during the study. The Principal Investigator can be contacted at [redacted]

By clicking the button below, you acknowledge:

- Your participation in the study is voluntary.
- You are over 18 years of age.
- You are aware that you may choose to terminate your participation at any time for any reason.

I consent, begin the study

I do not consent, I do not wish to participate

12:29

Choose the technology to be evaluated in this survey

- Solar Energy Source
- Wind Energy Source
- Electrolysis technology
- Liquid Hydrogen Storage
- Gaseous Hydrogen Storage
- Liquid Hydrogen Transportation
- Gaseous Hydrogen Transportation
- Refuelling Stations
- Other

← →

Powered by Qualtrics

12:29

Capacity (Electrolysis technology)

Utilisation: the degree to which the Electrolysis technology will be utilised is...

○ Very high
○ High
○ Medium
○ Low
○ Very low
○ NA

Intermittency: the degree to which the Electrolysis technology will lack constant levels of productivity will be...

Size: the degree to which the hydrogen volume produced or delivered from Electrolysis technology can be modified at the facility will be...

Comments on Capacity

Fig. 6. Survey features (authors' construction in Qualtrics software).

panel group of experts including only "twelve volunteer graduate and faculty researchers working in the hydrogen pathways program at the Institute of Transportation Studies at the University of California, Davis". McCarthy et al. (2007) "conducted the assessment as part of a 3-hour facilitated exercise, in which the panel was walked through the purpose and scope of the method, and the evaluation process". Although McCarthy's data collection process was very effective regarding the time demand and interaction of members, yet there might have been bias in the expert's judgment.

In this work, the target population consisted of experts associated with the five nodes of the HSC who were identified online and contacted via email or LinkedIn. Bias was treated by applying some specific rules. It seemed important to gather the opinions of people from different countries with considerable experience (more than five years). No country was excluded from the study (nine countries participated). The required experience did not necessarily relate to hydrogen but concerned the technology related to the HSC. For example, an expert on solar PV could provide feedback on the latter technology from the perspective of 2030. At least three experts were required per node to reach data saturation and consider the evaluation valid. Stakeholders in four areas were identified: industry, academia, consultancy, and regulatory agencies.

3.2.1. *Individual interviews and surveys (round 1)*

The experts were invited to participate in the study, as displayed in Fig. 5. Once the invitation was accepted, one or two research team members met with the expert and started a structured interview via Zoom or Teams. The second task was to fill in the survey to evaluate twenty-four metrics (Table 3). For this purpose, a generic questionnaire was made available in Qualtrics software (https://www.qualtrics.com). The expert could choose the technology of expertise (Fig. 6). The questions had three components: (1) "*In 2030*" + (2) definition of the metric (in Table 3) + (3) "*will be*" (e.g., "*In 2030, the degree to which the electrolysis technology will be utilized is expected is expected to be…* "). The expert could then answer, "Very low," "Low," "Medium," "High," "Very High," or "N/A," ask for clarification, and/or add comments.

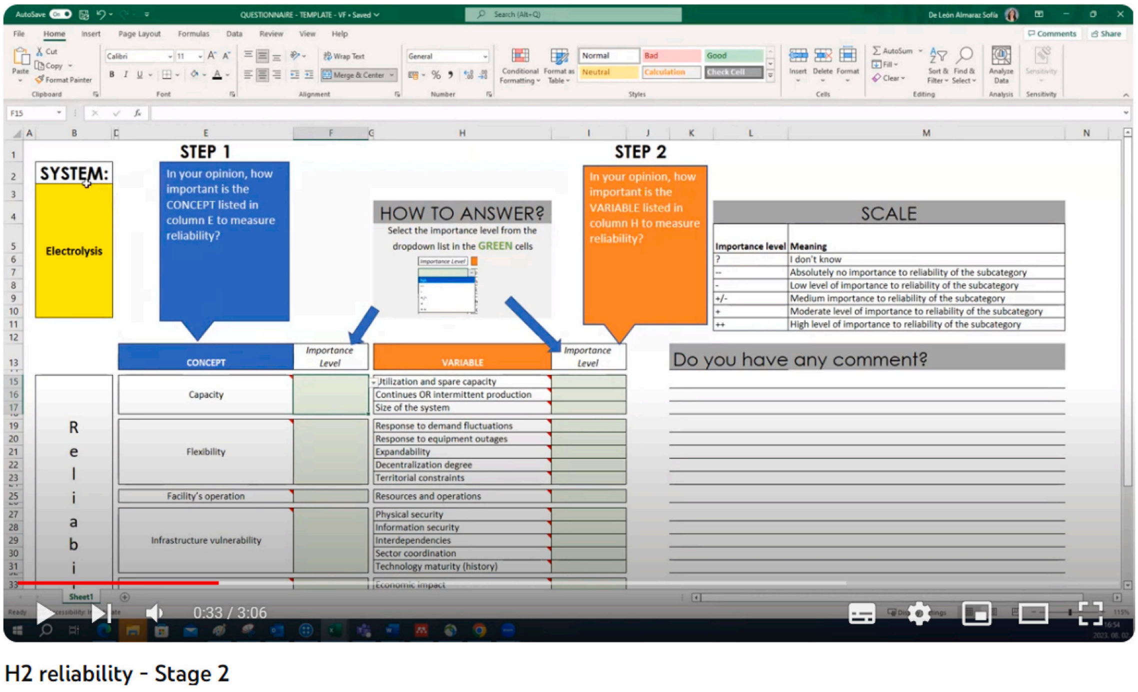
The questionnaire contained questions graded in two ways – for some items, like capacity, selecting "very high" represented high reliability and low risk (value of 1). In contrast, with others, such as intermittency, the scale was inverted, so "very high" represented low reliability and high risk (value of 5). It was decided to maintain this particularity to maintain the attention of the experts since, in the survey trial, it was noticed that this enhanced the comments shared about each item. A research team member could provide clarifications if needed and verify that the criteria under consideration were harmonized with those appraised by other participants. The team employed a guideline with examples for each metric and each Likert option used during the session, depending on the requirements. The estimated duration to completion was 45–60 min.

Opinions about metrics related to energy security (Table 3), applicable only to solar and wind energy, were not gathered directly through the survey because it was noticed in the survey trial that the experts did not have specific information about them (due to the international context). Instead, they were sourced from reports and validated by only one expert in the format of a potential future case study. A total of 24 metrics were assessed for the energy source node, while 18 metrics were assessed for each of the other nodes. The output from the individual evaluation was reliability values (translated to quantitative values for $u_{i,j}$), which are considered value added information. Qualitative analysis of the interview is outside the scope of this paper. Calculations and preparation for the next steps were completed (Fig. 5: "Data Analysis 1" and Section 3.1).

3.2.2. *e-Delphi method (rounds 2 and 3)*

In rounds 2–3 (Fig. 5), the evaluation centered on the importance ratings ($w_{i,j}$) of each metric, concept, and GO to assess reliability. Since a long list of variables was evaluated, it was considered it worth running a separate session for this to avoid mixing terms in Round 1 (purely dedicated to the reliability assessment). At this stage, at least two experts per technology were needed (minimum of 10 evaluations).

Due to time and location constraints, gathering participants together for the project's duration was impossible. For this reason, Fig. 5 shows two asynchronous e-Delphi rounds that used an Excel questionnaire to collect data for $w_{i,j.}$ The instructions for filling in the questionnaire were provided in a free video sharing website, and the questionnaire extension involved a page that improved the experts' understanding of the



STEP 1

SYSTEM:

Electrolysis

In your opinion, how important is the CONCEPT listed in column E to measure reliability?

HOW TO ANSWER?

Select the importance level from the dropdown list in the GREEN cells

STEP 2

In your opinion, how important is the VARIABLE listed in column H to measure reliability?

SCALE

Importance level	Meaning
?	I don't know
--	Absolutely no importance to reliability of the subcategory
-	Low level of importance to reliability of the subcategory
+/-	Medium importance to reliability of the subcategory
+	Moderate level of importance to reliability of the subcategory
++	High level of importance to reliability of the subcategory

	CONCEPT	Importance Level	VARIABLE	Importance Level
R e l i a b i	Capacity		Utilization and spare capacity	
			Continues OR intermittent production	
			Size of the system	
	Flexibility		Response to demand fluctuations	
			Response to equipment outages	
			Expandability	
			Decentralization degree	
			Territorial constraints	
	Facility's operation		Resources and operations	
	Infrastructure vulnerability		Physical security	
			Information security	
			Interdependencies	
			Sector coordination	
			Technology maturity (history)	
			Economic impact	

Do you have any comment?

H2 reliability - Stage 2

Fig. 7. Instructions and questionnaire used to rate the importance of metrics and concepts (authors' creation).

relationships between the variables with the brief description of importance levels scale (Fig. 7). Estimated time to complete was 10–15 min. Although this exercise cannot be considered a full example of the Delphi method because a brainstorming session or a panel discussion is not included, the e-Delphi method has been used in medical applications with acceptable results (e.g., Hall et al., 2019); one major advantage of this strategy is that it involves an element of transparency, albeit without influencing other experts' answers.

Based on the interview inputs, 20 experts were invited to participate, and 10 accepted. Importance was rated as "++" (high level of importance) to "–" (absolutely no importance), which was transformed on a scale from 1 to 5 for low importance to high importance, respectively. Answers were compared (Fig. 5: "Data analysis 2"), and if agreement was not reached (measure deviation >1 point on the Likert scale from 1 to 5), in Round 3 of the e-Delphi, a discussion took place. The research team shared the answers from the expert pairs to encourage rethinking, changes of opinion, or keeping the original answer and providing additional comments. This process allowed the participants to learn and think about the future, knowing others' opinions. An acceptance level regarding experts' agreement was not fixed since it was considered that the importance rate could be affected by diverse aspects when valid arguments were provided. For this reason, after Round 3, the experts' values were used to calculate the average importance rate per technology, as described in "Data Analysis 3" (Fig. 5); these inputs were used to make the final calculations for the reliability index.

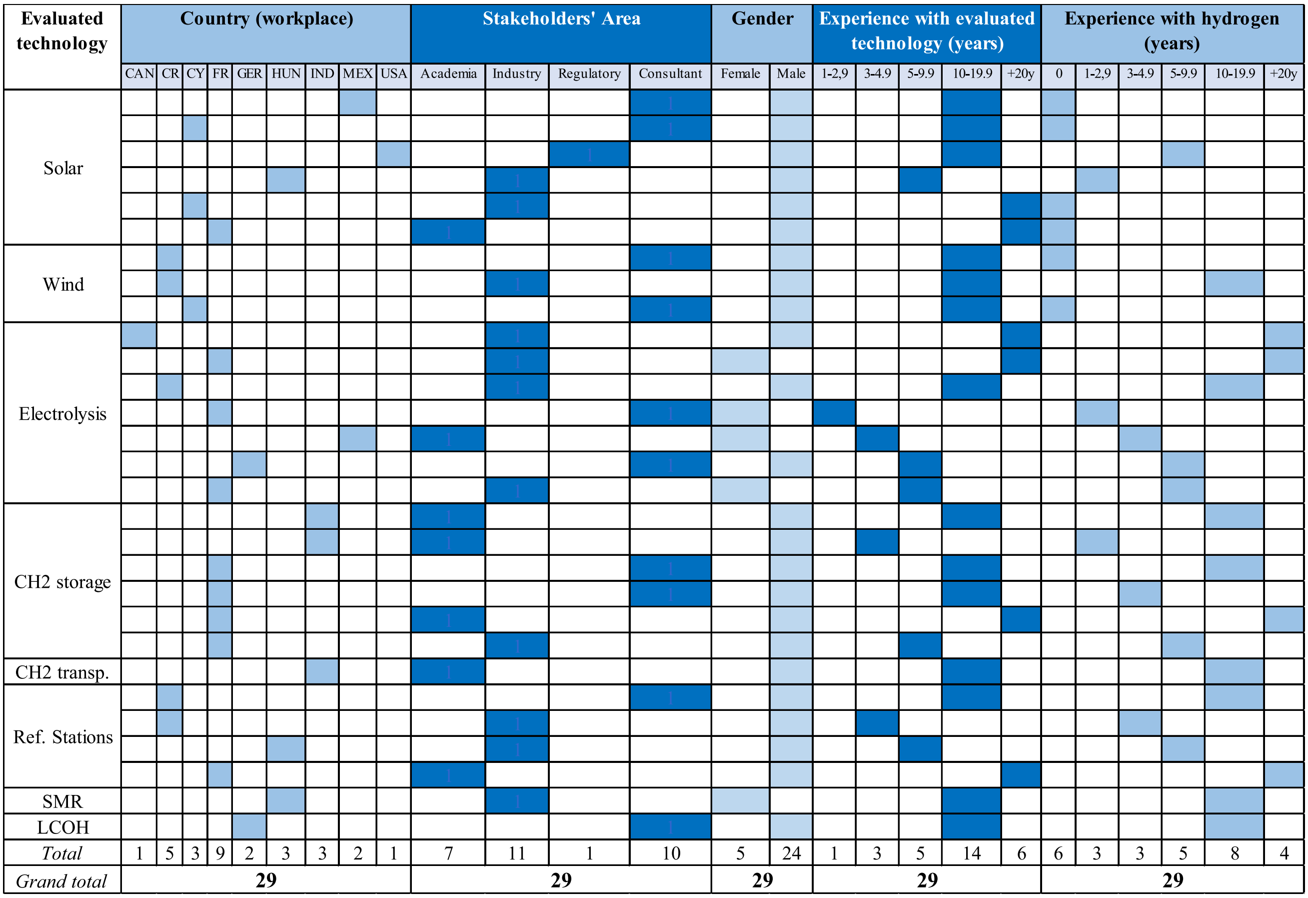
3.3. Case study

The "H_2 reliability project" started on 27 October 2022 and was completed on 18 Oct 2023. The project's objective was to find a way to quantify the reliability of HSCs. As previously explained, five researchers contributed to the development of the different stages of the project. The problem was delineated as international in scope. Five technologies (nodes) are part of supply chain options. HSC1 includes the following technologies: solar energy source, electrolysis, compressed hydrogen (CH_2) storage, CH_2 transportation, and refueling stations. HSC2 involves the same technologies as HSC1, except for the energy source specified as wind energy.

This international study acknowledges that stakeholders' opinions and perceptions may be influenced by their specific cultural contexts. Nevertheless, the defined scope of the problem accommodates broader regional or local perspectives. Experts were from America (31 %), Europe (58 %), and Asia (10 %). Therefore, the methodology emphasizes the involvement of diverse actors to ensure a comprehensive representation of various behaviors and motivations. Engaging a wide array of stakeholders is crucial for effectively mapping and evaluating these complex systems (Király et al., 2016).

Experts from academia (24 %), consultancy (34 %), industry (37 %), and regulatory bodies (1 %) were included to evaluate specific technologies (one or a maximum of two in separate sessions). With regard to the experience of experts, explicit experience with hydrogen was not required. Five years of experience was initially defined (58 % of the experts). However, for some technologies like electrolysis, some people had experience of less than three years (working on real implementation projects), so this constraint was relaxed to facilitate analysis of the potential correlation between the duration of experience and the ratings (in a second study).

In total, 89 experts were invited to participate in the study: 29 agreed to participate in an interview, 28 in the survey, and 27 allowed us to record the interviews. Only six interviews were face-to-face or hybrid. The duration of individual meetings with the experts ranged from 40 min to 1.5 h (including interview and survey).



Evaluated technology	Country (workplace)									Stakeholders' Area				Gender		Experience with evaluated technology (years)					Experience with hydrogen (years)					
	CAN	CR	CY	FR	GER	HUN	IND	MEX	USA	Academia	Industry	Regulatory	Consultant	Female	Male	1-2,9	3-4.9	5-9.9	10-19.9	+20y	0	1-2,9	3-4.9	5-9.9	10-19.9	+20y
Solar								■					1		■				■		■					
			■										1		■				■		■					
									■			1			■				■					■		
						■					1				■			■				■				
			■								1				■					■	■					
				■						1					■					■	■					
Wind		■											1		■				■		■					
		■									1				■				■						■	
			■										1		■				■		■					
Electrolysis	■										1				■					■						■
				■							1			■						■						■
		■									1				■				■						■	
				■									1	■		■						■				
								■		1				■			■						■			
					■								1		■			■						■		
				■							1			■				■						■		
CH2 storage							■			1					■				■						■	
							■			1					■		■					■				
				■									1		■				■						■	
				■									1		■				■				■			
				■						1					■					■						■
				■							1				■			■						■		
CH2 transp.							■			1					■				■						■	
Ref. Stations		■											1		■				■						■	
		■									1				■		■						■			
						■					1				■			■						■		
				■						1					■					■						■
SMR						■					1			■					■						■	
LCOH					■								1		■				■						■	
Total	1	5	3	9	2	3	3	2	1	7	11	1	10	5	24	1	3	5	14	6	6	3	3	5	8	4
Grand total	**29**									**29**				**29**		**29**					**29**					

Fig. 8. Profiles of experts participating in the study "H_2 reliability" (authors' creation).

The experts' profiles are depicted in Fig. 8 (including the available demographic details of target respondents). Initially, data saturation was targeted by recruiting a sufficient number of experts per node. However, due to time constraints, only one expert specializing in the transportation of compressed hydrogen agreed to participate. Unfortunately, experts on liquid hydrogen transportation and storage could not be enlisted.

Conversely, there was robust participation from experts in electrolysis (7), gaseous storage (6), solar energy (6), and refueling stations (4). Although only three experts were available for wind energy, this was deemed sufficient for data saturation in this node, warranting its inclusion. Additionally, experts contributed insights on other technologies, including Liquid Organic Hydrogen Carriers (LOHC) and Steam Methane Reforming (SMR).

In summary, technologies with fewer than three evaluations were excluded from the reliability calculation. However, transportation of gaseous hydrogen via tube trailer, despite having fewer than three evaluations, was identified for further analysis.

4. Results and discussion

4.1. Survey results

The quantitative results displayed in the box charts in Fig. 9 reveal the variability per metric for each technology prior to normalization, weighting, and aggregation. The scores are on a scale from 1 to 5, where values above 2.5 are generally indicative of lower reliability compared to scores below 2.5.

The methodology employed in this research provides a visual representation of the disparities in expert expectations regarding the assessed technologies in 2030. For instance, there is more agreement in terms of the reliability of electrolysis (including resources and operations). At the same time, conflicting viewpoints emerged concerning gaseous hydrogen storage, particularly in relation to the economic impact of the disruption of such technology. It is worth emphasizing that, at this level of analysis, it is possible to begin to discern the underlying reasons for these differing perspectives.

Upon aggregating the variable results, even before normalizing and

Fig. 9. Survey findings per technology categorized by reliability metrics (prior to normalization) (authors' creation).

accounting for the importance weights, a medium level of reliability was expected for all technologies. The process involved using 18 metrics in calculating utility (u) for electrolysis, CH2 storage, CH2 transportation, and refueling stations and 24 metrics for energy sources. In this context, the reliability for all technologies ranged between 2.40 and 2.66, classifying them as having medium reliability. Notably, refueling stations emerged as the best-rated technology (before normalization), whereas gaseous transportation was considered the least reliable (based only on one expert's judgment). Utilizing the survey and interview-based methodology theoretically enables a qualitative analysis of each variable per technology, although this is beyond the scope of this paper. To compare these outcomes with a reference, the TRL values reported by the IEA and shown in Table 1 can be used. In this framework, one metric is technology maturity, and it is evident that gaseous storage is expected to be the most mature technology both presently and in 2030, with a TRL of 11. PV and HRS results follow a similar trend to the one reported in Table 1. There is less agreement about wind energy and electrolysis, with more optimism about the current readiness level than the maturity expected in 2030 in an HSC. The variability in Fig. 9 displays the opinions of experts in evaluating the variables associated with specific metrics.

4.2. MAUT results

Multi-attribute utility results were derived from the average of the aggregated utility scores obtained from the expert survey, which was conducted on a per-technology basis rather than per pathway as in McCarthy et al. (2007). The importance ratings were assigned by a smaller group of experts, as elaborated in Section 3.2.2. Fig. 10 provides the average ratings and aggregated utility and importance scores for each technology and supply chain, encompassing scores for all the metrics, concepts, GOs (adequacy and security), and the reliability index per hydrogen supply chain (HSC1 and HSC2). It is important to note that the results and conclusions presented in this work are preliminary and cannot be generalized. However, they serve the purpose of testing the methodology and identifying trends in expectations for 2030.

The comparison of the two green H_2 supply chains is given in Fig. 10. HSC1 used a solar energy source, whereas HSC2 employed wind energy, setting them apart. The result suggests that experts assessed wind and solar energy sources similarly, as evidenced by the utility values in Fig. 9. However, there are some differences in the importance rating of metrics used to evaluate the reliability of the two options. Consequently, there is a slight discrepancy before and after scaling, weighting, and aggregating the different technologies (e.g., the wind reliability assessment was better than the one for solar before including the importance rating). Notably, the results are marginally more favorable for solar (adequacy: 0.225, security: 0.257) than wind (adequacy: 0.246, security: 0.245), as depicted in Fig. 11.

HSC1 (solar) is anticipated to be more reliable in terms of both adequacy (1.263) and security (1.248), resulting in a reliability index of 2.51 (Fig. 11). Conversely, HSC2, the wind pathway, yielded aggregated scores of 1.285 and 1.236 for adequacy and security, culminating in a final reliability score of 2.52 (recall that in the proposed model, higher utility [“u”] scores represent lower reliability). The slight difference between these scores can be regarded as negligible, indicating that both technologies offer a similar level of reliability, consistent with the medium reliability expected for HSC1 and HSC2 by 2030.

Fig. 11 illustrates the scale used for each calculation in relation to the aggregation steps.

Utility scores can be compared at various levels to assess the relative strengths and weaknesses of each HSC. Given that the sole distinction between HSC1 and HSC2 is the type of energy source, it is relatively straightforward to identify variations in broader categories, particularly regarding flexibility and the consequences of technology disruption. Notably, the importance awarded to the expert ratings strongly influences the outcome. For instance, the reliability ranking of technologies can undergo significant shifts, as exemplified by solar energy (Figs. 9 and 10) transitioning from 0.5 (in the simple reliability evaluation from the survey) to a relatively higher reliability index (0.48) following the incorporation of importance weights. Moreover, factors like utilization of resources and operations are considered less critical with wind energy compared to solar. In contrast, aspects such as response to equipment outages, expandability, or impact on interdependent systems are more significant for wind than solar. These

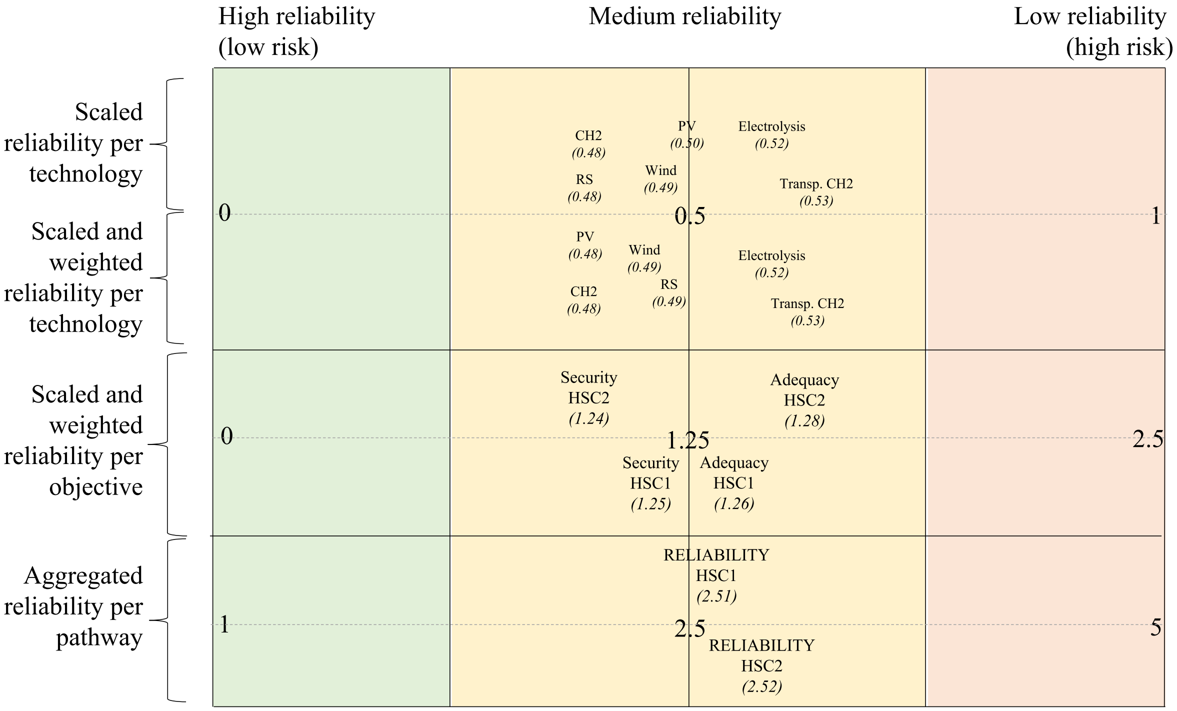
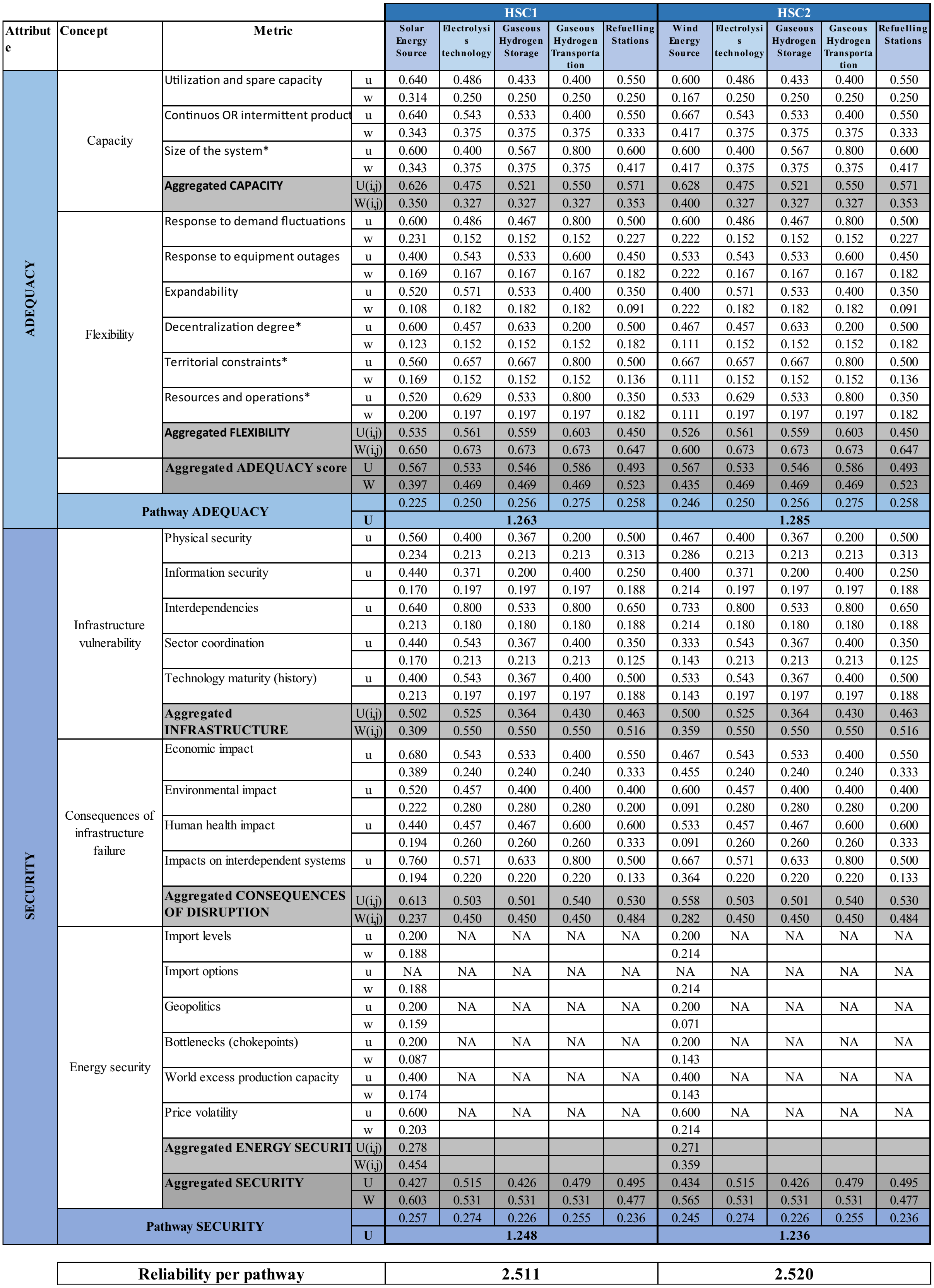


Fig. 10. Reliability and risk results per technology and pathway before and after aggregation (authors' creation).



				HSC1					HSC2				
Attribute	**Concept**	**Metric**		**Solar Energy Source**	**Electrolysis technology**	**Gaseous Hydrogen Storage**	**Gaseous Hydrogen Transportation**	**Refuelling Stations**	**Wind Energy Source**	**Electrolysis technology**	**Gaseous Hydrogen Storage**	**Gaseous Hydrogen Transportation**	**Refuelling Stations**
ADEQUACY	Capacity	Utilization and spare capacity	u	0.640	0.486	0.433	0.400	0.550	0.600	0.486	0.433	0.400	0.550
			w	0.314	0.250	0.250	0.250	0.250	0.167	0.250	0.250	0.250	0.250
		Continuos OR intermittent product	u	0.640	0.543	0.533	0.400	0.550	0.667	0.543	0.533	0.400	0.550
			w	0.343	0.375	0.375	0.375	0.333	0.417	0.375	0.375	0.375	0.333
		Size of the system*	u	0.600	0.400	0.567	0.800	0.600	0.600	0.400	0.567	0.800	0.600
			w	0.343	0.375	0.375	0.375	0.417	0.417	0.375	0.375	0.375	0.417
		Aggregated CAPACITY	U(i,j)	0.626	0.475	0.521	0.550	0.571	0.628	0.475	0.521	0.550	0.571
			W(i,j)	0.350	0.327	0.327	0.327	0.353	0.400	0.327	0.327	0.327	0.353
	Flexibility	Response to demand fluctuations	u	0.600	0.486	0.467	0.800	0.500	0.600	0.486	0.467	0.800	0.500
			w	0.231	0.152	0.152	0.152	0.227	0.222	0.152	0.152	0.152	0.227
		Response to equipment outages	u	0.400	0.543	0.533	0.600	0.450	0.533	0.543	0.533	0.600	0.450
			w	0.169	0.167	0.167	0.167	0.182	0.222	0.167	0.167	0.167	0.182
		Expandability	u	0.520	0.571	0.533	0.400	0.350	0.400	0.571	0.533	0.400	0.350
			w	0.108	0.182	0.182	0.182	0.091	0.222	0.182	0.182	0.182	0.091
		Decentralization degree*	u	0.600	0.457	0.633	0.200	0.500	0.467	0.457	0.633	0.200	0.500
			w	0.123	0.152	0.152	0.152	0.182	0.111	0.152	0.152	0.152	0.182
		Territorial constraints*	u	0.560	0.657	0.667	0.800	0.500	0.667	0.657	0.667	0.800	0.500
			w	0.169	0.152	0.152	0.152	0.136	0.111	0.152	0.152	0.152	0.136
		Resources and operations*	u	0.520	0.629	0.533	0.800	0.350	0.533	0.629	0.533	0.800	0.350
			w	0.200	0.197	0.197	0.197	0.182	0.111	0.197	0.197	0.197	0.182
		Aggregated FLEXIBILITY	U(i,j)	0.535	0.561	0.559	0.603	0.450	0.526	0.561	0.559	0.603	0.450
			W(i,j)	0.650	0.673	0.673	0.673	0.647	0.600	0.673	0.673	0.673	0.647
		Aggregated ADEQUACY score	U	0.567	0.533	0.546	0.586	0.493	0.567	0.533	0.546	0.586	0.493
			W	0.397	0.469	0.469	0.469	0.523	0.435	0.469	0.469	0.469	0.523
	Pathway ADEQUACY			0.225	0.250	0.256	0.275	0.258	0.246	0.250	0.256	0.275	0.258
			U	**1.263**					**1.285**				
SECURITY	Infrastructure vulnerability	Physical security	u	0.560	0.400	0.367	0.200	0.500	0.467	0.400	0.367	0.200	0.500
				0.234	0.213	0.213	0.213	0.313	0.286	0.213	0.213	0.213	0.313
		Information security	u	0.440	0.371	0.200	0.400	0.250	0.400	0.371	0.200	0.400	0.250
				0.170	0.197	0.197	0.197	0.188	0.214	0.197	0.197	0.197	0.188
		Interdependencies	u	0.640	0.800	0.533	0.800	0.650	0.733	0.800	0.533	0.800	0.650
				0.213	0.180	0.180	0.180	0.188	0.214	0.180	0.180	0.180	0.188
		Sector coordination	u	0.440	0.543	0.367	0.400	0.350	0.333	0.543	0.367	0.400	0.350
				0.170	0.213	0.213	0.213	0.125	0.143	0.213	0.213	0.213	0.125
		Technology maturity (history)	u	0.400	0.543	0.367	0.400	0.500	0.533	0.543	0.367	0.400	0.500
				0.213	0.197	0.197	0.197	0.188	0.143	0.197	0.197	0.197	0.188
		Aggregated INFRASTRUCTURE	U(i,j)	0.502	0.525	0.364	0.430	0.463	0.500	0.525	0.364	0.430	0.463
			W(i,j)	0.309	0.550	0.550	0.550	0.516	0.359	0.550	0.550	0.550	0.516
	Consequences of infrastructure failure	Economic impact	u	0.680	0.543	0.533	0.400	0.550	0.467	0.543	0.533	0.400	0.550
				0.389	0.240	0.240	0.240	0.333	0.455	0.240	0.240	0.240	0.333
		Environmental impact	u	0.520	0.457	0.400	0.400	0.400	0.600	0.457	0.400	0.400	0.400
				0.222	0.280	0.280	0.280	0.200	0.091	0.280	0.280	0.280	0.200
		Human health impact	u	0.440	0.457	0.467	0.600	0.600	0.533	0.457	0.467	0.600	0.600
				0.194	0.260	0.260	0.260	0.333	0.091	0.260	0.260	0.260	0.333
		Impacts on interdependent systems	u	0.760	0.571	0.633	0.800	0.500	0.667	0.571	0.633	0.800	0.500
				0.194	0.220	0.220	0.220	0.133	0.364	0.220	0.220	0.220	0.133
		Aggregated CONSEQUENCES OF DISRUPTION	U(i,j)	0.613	0.503	0.501	0.540	0.530	0.558	0.503	0.501	0.540	0.530
			W(i,j)	0.237	0.450	0.450	0.450	0.484	0.282	0.450	0.450	0.450	0.484
	Energy security	Import levels	u	0.200	NA	NA	NA	NA	0.200	NA	NA	NA	NA
			w	0.188					0.214				
		Import options	u	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
			w	0.188					0.214				
		Geopolitics	u	0.200	NA	NA	NA	NA	0.200	NA	NA	NA	NA
			w	0.159					0.071				
		Bottlenecks (chokepoints)	u	0.200	NA	NA	NA	NA	0.200	NA	NA	NA	NA
			w	0.087					0.143				
		World excess production capacity	u	0.400	NA	NA	NA	NA	0.400	NA	NA	NA	NA
			w	0.174					0.143				
		Price volatility	u	0.600	NA	NA	NA	NA	0.600	NA	NA	NA	NA
			w	0.203					0.214				
		Aggregated ENERGY SECURIT	U(i,j)	0.278					0.271				
			W(i,j)	0.454					0.359				
		Aggregated SECURITY	U	0.427	0.515	0.426	0.479	0.495	0.434	0.515	0.426	0.479	0.495
			W	0.603	0.531	0.531	0.531	0.477	0.565	0.531	0.531	0.531	0.477
	Pathway SECURITY			0.257	0.274	0.226	0.255	0.236	0.245	0.274	0.226	0.255	0.236
			U	**1.248**					**1.236**				
	Reliability per pathway			**2.511**					**2.520**				

Fig. 11. Reliability index results for HSC1 and HSC2 (* new metrics) (authors' creation).

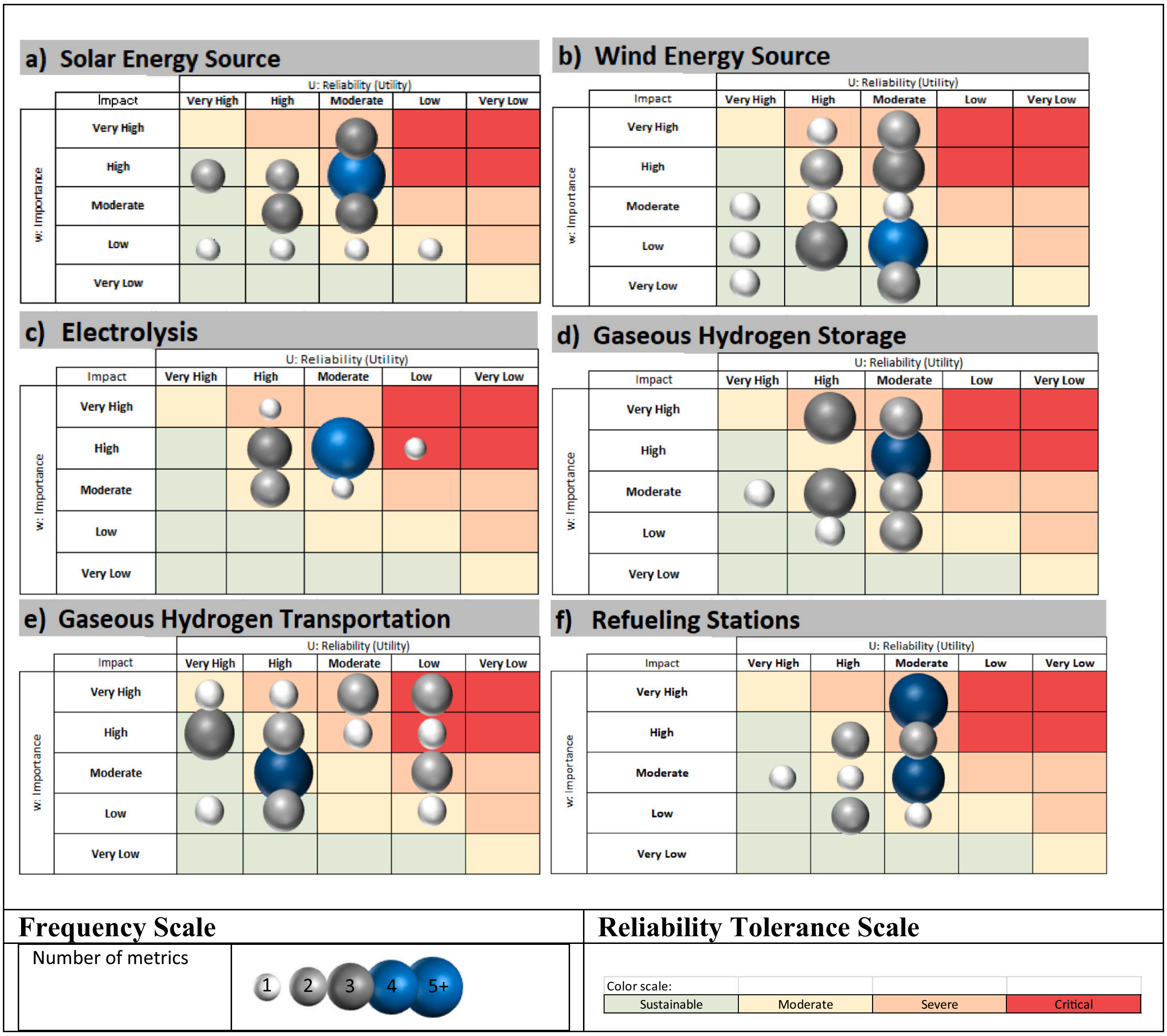


Fig. 12. Reliability heat maps for hydrogen supply chain technologies.

variations reflect expert assessments, which consider the system as a whole and evaluate the impact of the failure of independent components, such as a single solar panel or turbine, as indicated during the interview and survey sessions.

4.3. Reliability matrix per technology (risk binning matrix)

The green hydrogen supply chain projected by 2030 was analyzed using heat maps, where the metrics are ranked based on their importance and reliability (Fig. 12). The risk assessment leads to a shared visualization of the quantitative data and suggests prioritizing the reliability rating of each supply chain node.

Here are the key highlights for each selected technology.

For the solar energy source (Fig. 12a), severe risk perceptions are associated with the following metrics: Resources and operation, Territorial conditions to install, Physical security, Economic impact, Price Volatility, Response to demand fluctuations, Size of the system, Intermittency, and Utilization.

For the wind energy source (Fig. 12b), severe risk perceptions are associated with the following metrics: Impacts on interdependent system, Economic impact, Interdependencies, Intermittency, Response to demand fluctuations, Response to equipment outages, Size of the system.

Electrolysis (Fig. 12c) is associated with severe risk perceptions for the following metrics: Impacts on interdependent systems, Environmental impact, Economic impact, Technology maturity (history), Sector Coordination, Resources and operation, Expandability, and Response to equipment outages. Furthermore, the metric Interdependencies is perceived as critical.

For Gaseous Hydrogen Storage (Fig. 12d), severe risk perceptions are associated with the following metrics: Human health impact, Economic impact, Physical security, Resources and operation, Decentralization, Response to equipment outages, Response to demand fluctuations, Size, and Intermittency.

For Gaseous Hydrogen Transportation (Fig. 12e), critical risk is perceived for the following metrics: Resources and operation, Response to demand fluctuations, and Size. Furthermore, Fig. 12e indicates the severe risk perception related to the following concepts: Human health impact, Economic impact, Interdependencies, Territorial constraints, and Response to equipment outages.

With Refueling stations (Fig. 12f) severe risk perceptions are associated with the following metrics: Human health impact, Economic impact, Physical security, Decentralization, Response to demand fluctuations, Size, and Intermittency.

In summary, the main critical reliability ratings are associated with the following:

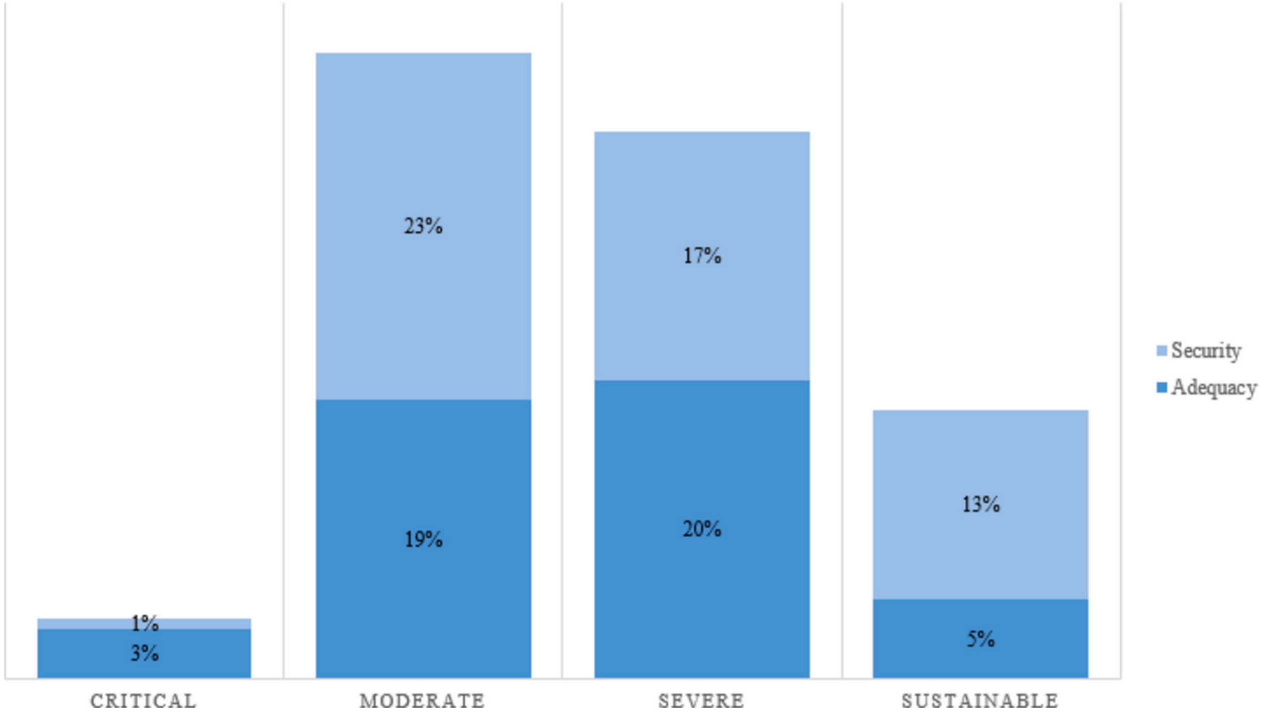


Fig. 13. Perception of cumulative risks associated with the green hydrogen supply chain for the two main components of reliability (adequacy and security).

- Electrolysis: interdependencies (concept: infrastructure; GO: security).
- Gaseous hydrogen transportation: resources and operation, response to demand fluctuations (concept: flexibility; GO: adequacy), and size (concept: capacity; GO: adequacy).

Finally, considering the technologies' reliability performance, 53 % of the metrics are related to the Security category, associated with just 1 % critical risk and 17 % severe risk ratings. For the Adequacy category, 47 % of the concepts attract a 3 % critical risk rating and 20 % a severe risk rating (Fig. 13).

5. Conclusion

This study was designed to appraise the reliability of green hydrogen supply chains by 2030. Given the futuristic nature of the evaluation and ongoing technological developments in hydrogen-related fields, conventional risk assessment methods were deemed unsuitable. The model of McCarthy et al.' (2007) was initially adopted, enhancing it by integrating new metrics and a novel data collection methodology. In summary, both qualitative (HS) and quantitative (MAUT) approaches are proposed to assess the reliability of prospective HSCs.

The findings suggest that multi-attribute utility theory may be appropriate for strategically evaluating hydrogen supply chains, provided the metrics are well-defined. However, these results may not be generalizable, as they rely on subjective input. In the international case study conducted; horizon scanning proved valuable for gathering experts' insights on the progress of hydrogen technologies in their respective regions. The chosen time horizon was generally considered adequate and realistic by 96 % of the experts. Twenty-nine experts from nine countries evaluated the reliability of technologies related to green hydrogen supply chains (solar and wind energy, electrolysis, gaseous storage, and transportation and refueling stations). While individual evaluations minimize bias, experts' willingness to collaborate with counterparts from other regions was evident. The quantitative reliability analysis suggests that hydrogen experts are concerned about the projected reliability of green HSC.

Experts project moderate reliability for green HSCs by 2030. Among these, solar HSCs are considered slightly more reliable than wind HSCs (aggregated results indicated a minimal disparity in reliability projections for both pathways). Critical risks were identified for electrolysis and transportation technologies. For electrolysis, the main obstacle is related to its interdependence with other systems, primarily renewable energy sources. For transportation, the critical points are associated with resources, response to demand fluctuations, and size. It is highlighted that only one expert was ready to provide feedback on the transportation of gaseous hydrogen; therefore, this area of HSC requires deeper analysis.

Rating the reliability index through the survey may be simplified when more input about the importance of the metrics and concepts is available and statistical analysis is performed. So far, the proposed methodology allowed the calculation of the reliability index per technology and supply chain, revealing variations in expert opinions and highlighting critical reliability metrics for technologies.

Applying a methodology like the one proposed for the “H_2 reliability project” necessitates the involvement of a multidisciplinary team with quantitative and qualitative competencies. It is essential to note that the results presented in this paper exclusively pertain to the survey application used to estimate reliability. Qualitative analysis of the interviews is currently in progress, associated with specific comments addressing reliability metrics for individual technologies. After completing the qualitative analysis, a more in-depth examination of how the former influence decision-making among HSC stakeholders can be provided.

5.1. Key contributions

The proposed methodology uses engineering and social science insights to answer questions related to the future energy reliability of green HSCs. The theoretical contributions are (1) identification of the state of the art related to the reliability of prospective HSCs; (2) identification of research gaps, as presented in Section 2.2; (3) a critical evaluation of McCarthy et al.'s (2007) approach, and the proposal of additional reliability metrics; (4) the exploration and proposal of qualitative methods applicable to future studies and selection of horizon scanning to investigate technological expectations concerning HSC's reliability in 2030; (5) the conceptualization of a mixed methods framework which includes an extended version of McCarthy et al.'s (2007) reliability model using MAUT coupled with HS to design the data collection instruments (survey and interview questionnaires).

The methodological contributions are as follows: (1) a reliability analysis at various levels: The case study applied to the macro level of HSC, and the methodology addresses the unique challenges posed by new infrastructure, including interdependencies and their impact. (2) Facilitation of a deep exploration of specific metrics and variables, highlighting the importance of preserving raw data before aggregation, thus mitigating the risk of misinterpretation and overly generalized conclusions. Four new metrics were added to the proposed framework and analysis: size, territorial conditions, decentralization degree, and resources. (3) The methodology allows for the customized importance rating evaluation of proposed metrics representing the reliability of different technologies, offering flexibility and promoting consensus-seeking via the e-Delphi method. (4) The multidisciplinary and international approach has the flexibility to incorporate diverse viewpoints from various contexts and backgrounds, and the prospective analysis enables the identification of potential future risks. (5) Employing mixed methods allows for a holistic reliability assessment and examination of both numerical data and qualitative insights. Quantitative analysis was employed for this paper, while qualitative analysis is in progress. (6) The methodology facilitates the identification of key takeaways by presenting results through a range of graphical representations, such as plot charts and risk matrix-heat maps, enabling a clear understanding of hotspots.

The proposed methodology may be useful for hydrogen stakeholders, including government, industry, researchers, reliability experts, and society in general. It may also have implications regarding the systemic design of future HSCs that can reliably serve society.

5.2. Challenges and limitations

Several challenges were identified in the initial application of the methodology, including those related to language, survey length and complexity, and the deductive nature of structured interviews. Support with most interviews and completing the survey was necessary for validation purposes. Engaging in discussions with a diverse group of experts required a solid understanding of hydrogen technologies and the context to facilitate meaningful conversations and offer guidance and clarification throughout the survey process. Given the extensive number of survey items and potential response options, this support was indispensable in ensuring accurate and consistent responses.

McCarthy et al. (2007) evaluated pathways that integrated both centralized and distributed (decentralized) systems. The experts in that study perceived distributed hydrogen production and on-site utilization at refueling stations as advantageous. Suitability may be enhanced by increasing the flexibility to adapt to fluctuations in demand in terms of volume and geographical distribution. Due to the study's interest in green hydrogen, the analyzed hydrogen supply chains did not incorporate large centralized production options. Furthermore, an analysis of the transport and storage of liquid hydrogen or gas using pipelines was not included due to the absence of participants with expertise in these technologies. The outcomes underscore the critical importance of evaluating HSC incorporating SMR, now a widely used technology.[2] Such an evaluation could serve as a baseline, facilitating the comparison of new potential pathways against the status quo regarding reliability instead of using only TRLs. The suggested method is a more complex evaluation than just using a 1–5 scale, as shown in Fig. 10. It allows for a full examination of HSCs' reliability compared to established standards.

The academic research encountered difficulties concerning contacting various groups of people, primarily due to time and location constraints that necessitated the division of data collection. This situation led us to use the e-Delphi method. Although the heterogeneous profile of experts enriched the discussions, it also resulted in some ambiguity in defining specific contexts, given the global scope of this study. However, maintaining a broad definition of the technology system was a necessity. Attempting to delve into local or global contexts and confidently rate the metrics proved difficult for several experts. What this study calls a metric may indeed cover multiple specific metrics. In initiating the questionnaire using McCarthy's framework, this research team handpicked the metrics, concepts, and objectives based on specific criteria. Integrating these metrics into concepts remains a subjective process and necessitates validation. Further, making evaluations using the current instrument is complex as it involves appraising some metrics based on reliability and others based on risk. This differential approach required additional explanation and data treatment to ensure consistency. Creating a prospective study and defining the scale for each metric also demanded significant time and resources. The sample size could be expanded to include additional nodes for a more detailed analysis. Because the study is also semi-qualitative, the expertise and judgment of the participating experts and the researchers' biases introduce uncertainty about the framework definition and interpretations of results. Several strategies were used to mitigate these biases, highlighting the importance of multidisciplinary work.

6. Perspectives

Several additional perspectives can be considered. Creating a baseline reliability index from the current status quo of technologies would serve as a valuable reference point for comparison. Establishing low and high-reliability values could enhance comparison among technologies in relation to the status quo, such as SMR technology and associated supply chain technologies, which represent mature pathways. Conducting a sensitivity analysis on importance ratings would help gauge their impact on the final calculations. Expanding the pool of experts in the study would enhance the validity of the results and provide a broader spectrum of insights. Moreover, organizing panel discussions, an alternative methodology, might be beneficial for comparing the outcomes with individual assessments and exploring possibilities for agreement. For modeling and safety purposes, the analysis would benefit from including a technical risk assessment. Conducting statistical analysis on a larger sample will be crucial for evaluating the validity of the questionnaire and its diverse items. Since the approach considers the future, integrating financial reliability into the methodology would generate valuable insights into the economic aspects of reliability. Following this first empirical exercise, it would be valuable to implement the methodology into a specific case study focusing on current or near-future perspectives to examine a more practical context. Overall efficiency and reliability performance can also be determined by modeling the interdependencies between various parameters at each HSC node.

CRediT authorship contribution statement

Sofía De-León Almaraz: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Tchougoune Moustapha Mai:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Iris Rocio Melendez:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **M.K. Loganathan:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Catherine Azzaro-Pantel:** Writing – review & editing,

---

[2] One interview with an expert in Steam Methane Reforming (SMR) was conducted, who responded to the survey using the perspective of 2030.

Validation, Supervision, Resources, Methodology, Investigation, Conceptualization.

Acknowledgements

We would like to thank all the experts who accepted to participate in this study and the reviewers for their valuable comments. This project was supported by the National Research, Development and Innovation Fund of Hungary [grant number: OTKA FK 146575]. We are grateful to Dávid Losonci, Judit Gáspár, Gabriella Kiss, Moreno Frau and Grzegorz Leszczyński for their support and helpful comments and discussions. We also thank Simon Milton for proofreading.

Data availability

The data that has been used is confidential.

References

Al-Douri, A., Groth, K.M., 2024. Hydrogen production via electrolysis: state-of-the-art and research needs in risk and reliability analysis. Int. J. Hydrog. Energy 63, 775–785. https://doi.org/10.1016/j.ijhydene.2024.03.188.

Azzaro-Pantel, C., 2018. Hydrogen Supply Chains, 1st ed. Academic Press.

Bakker, S., Van Lente, H., Meeus, M., 2011. Arenas of expectations for hydrogen technologies. Technol. Forecast. Soc. Change 78, 152–162. https://doi.org/10.1016/j.techfore.2010.09.001.

Bakker, S., van Lente, H., Meeus, M.T.H., 2012. Credible expectations — the US Department of Energy's hydrogen program as enactor and selector of hydrogen technologies. Technol. Forecast. Soc. Change 79, 1059–1071. https://doi.org/10.1016/j.techfore.2011.09.007.

Bavaresco, M.V., D'Oca, S., Ghisi, E., Lamberts, R., 2020. Methods used in social sciences that suit energy research: a literature review on qualitative methods to assess the human dimension of energy use in buildings. Energy Build. 209, 109702. https://doi.org/10.1016/j.enbuild.2019.109702.

Blank, R., 2004. The Basics of Reliability. Productivity Press, New York.

Bonaccorsi, A., Apreda, R., Fantoni, G., 2020. Expert biases in technology foresight. Why they are a problem and how to mitigate them. Technol. Forecast. Soc. Chang. 151, 119855. https://doi.org/10.1016/j.techfore.2019.119855.

Borup, M., Brown, N., Konrad, K., Van Lente, H., 2006. The sociology of expectations in science and technology. Tech. Anal. Strat. Manag. 18, 285–298. https://doi.org/10.1080/09537320600777002.

Brosch, T., Patel, M.K., Sander, D., 2014. Affective influences on energy-related decisions and behaviors. Front. Energy Res. 2, 81544. https://doi.org/10.3389/FENRG.2014.00011/BIBTEX.

Brosch, T., Sander, D., Patel, M.K., 2016. Editorial: behavioral insights for a sustainable energy transition. Front. Energy Res. 4, 193310. https://doi.org/10.3389/FENRG.2016.00015/BIBTEX.

Budde, B., Alkemade, F., Weber, K.M., 2012. Expectations as a key to understanding actor strategies in the field of fuel cell and hydrogen vehicles. Technol. Forecast. Soc. Change 79, 1072–1083. https://doi.org/10.1016/j.techfore.2011.12.012.

Chauhan, A., Golestani, N., Yazdi, M., Njue, J.C.W., Abbassi, R., Salehi, F., 2023. A novel integrated methodology for human reliability assessment in hydrogen fuelling stations. Int. J. Hydrog. Energy 48, 12917–12933. https://doi.org/10.1016/j.ijhydene.2022.12.181.

Chi, L., Qadrdan, M., Chaudry, M., Su, H., Zhang, J., 2024. Reliability of net-zero energy systems for South Wales. Appl. Energy 369, 123583. https://doi.org/10.1016/j.apenergy.2024.123583.

Clean Hydrogen Partnership, 2023. Interested in Hydrogen Safety? [WWW Document]. URL. https://www.clean-hydrogen.europa.eu/media/news/interested-hydrogen-safety-2023-06-07_en. (Accessed 12 June 2023).

Correa-Jullian, C., Groth, K.M., 2022. Data requirements for improving the quantitative risk assessment of liquid hydrogen storage systems. Int. J. Hydrog. Energy 47, 4222–4235. https://doi.org/10.1016/j.ijhydene.2021.10.266.

De-León Almaraz, S., Azzaro-Pantel, C., Montastruc, L., Domenech, S., 2014. Hydrogen supply chain optimization for deployment scenarios in the Midi-Pyrénées region, France. Int. J. Hydrog. Energy 39, 11831–11845. https://doi.org/10.1016/j.ijhydene.2014.05.165.

De-León Almaraz, S., Rácz, V., Azzaro-Pantel, C., Szántó, Z.O., 2022. Multiobjective and social cost-benefit optimisation for a sustainable hydrogen supply chain: application to Hungary. Appl. Energy 325, 119882. https://doi.org/10.1016/J.APENERGY.2022.119882.

De-León Almaraz, S., Kocsis, T., Azzaro-Pantel, C., Szántó, Z.O., 2024. Identifying social aspects of the hydrogen economy: review, synthesis, and research perspectives. Int. J. Hydrog. Energy 49, 601–618. https://doi.org/10.1016/j.ijhydene.2023.10.043.

Ebeling, C.E., 1997. An Introduction to Reliability and Maintainability Engineering. McGraw Hill.

Ehrhart, B.D., Hecht, E., Groth, K., 2021. Hydrogen Risk Assessment Models (HyRAM) (V.3.1) (Technical Reference Manual). https://doi.org/10.2172/1784842.

Elshurafa, A.M., Muhsen, A.R., Felder, F.A., 2022. Cost, footprint, and reliability implications of deploying hydrogen in off-grid electric vehicle charging stations: a GIS-assisted study for Riyadh, Saudi Arabia. Int. J. Hydrog. Energy 47, 32641–32654. https://doi.org/10.1016/j.ijhydene.2022.07.160.

European Commission, 2020. National Energy and Climate Plans (NECPs). https://energy.ec.europa.eu/topics/energy-strategy/national-energy-and-climate-plans-necps_en.

European Commission, 2022. Communication From the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. REPowerEU Plan [WWW Document]. EUR-Lex. URL. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A230%3AFIN&qid=1653033742483. (Accessed 26 June 2022).

European Commission, n.d. Energy strategy | Energy [WWW Document]. URL https://ec.europa.eu/energy/topics/energy-strategy-and-energy-union_en (accessed 1.7.21).

Fazli-Khalaf, M., Naderi, B., Mohammadi, M., Pishvaee, M.S., 2020. Design of a sustainable and reliable hydrogen supply chain network under mixed uncertainties: a case study. Int. J. Hydrog. Energy 45 (59), 34503–34531. https://doi.org/10.1016/j.ijhydene.2020.05.276.

Fetanat, A., Tayebi, M., 2024. Sustainability and reliability-based hydrogen technologies prioritization for decarbonization in the oil refining industry: a decision support system under single-valued neutrosophic set. Int. J. Hydrog. Energy 52, 765–786. https://doi.org/10.1016/j.ijhydene.2023.08.229.

Fortes, P., Alvarenga, A., Seixas, J., Rodrigues, S., 2015. Long-term energy scenarios: bridging the gap between socio-economic storylines and energy modeling. Technol. Forecast. Soc. Change 91, 161–178. https://doi.org/10.1016/j.techfore.2014.02.006.

Franki, V., Višković, A., 2015. Energy security, policy and technology in South East Europe: presenting and applying an energy security index to Croatia. Energy 90, 494–507. https://doi.org/10.1016/J.ENERGY.2015.07.087.

Gáspár, J., Hideg, É., Köves, A., 2021. Future in the present: participatory futures research methods in economic higher education – the development of future competencies. J. Futures Stud. 26. https://doi.org/10.6531/JFS.202112_26(2).0001.

González Palencia, J.C.G., Itoi, Y., Araki, M., 2022. Design of a hydrogen production system considering energy consumption, water consumption, CO2 emissions and cost. Energies 15, 7938. https://doi.org/10.3390/EN15217938.

Groth, K.M., Al-Douri, A., 2023. Hydrogen safety, risk, and reliability analysis. In: Hydrogen Economy. Elsevier, pp. 487–510. https://doi.org/10.1016/B978-0-323-99514-6.00012-1.

Groth, K.M., Al-Douri, A., West, M., Hartmann, K., Saur, G., Buttner, W., 2024. Design and requirements of a hydrogen component reliability database (HyCReD). Int. J. Hydrog. Energy 51, 1023–1037. https://doi.org/10.1016/j.ijhydene.2023.07.165.

Guo, Q., Nojavan, S., Lei, S., Liang, X., 2021. Potential evaluation of power-to-hydrogen-to methane conversion technology in robust optimal energy management of a multi-energy industrial park. Int. J. Hydrog. Energy. https://doi.org/10.1016/J.IJHYDENE.2021.07.148.

Hall, D.A., Hibbert, A., Smith, H., Haider, H.F., Londero, A., Mazurek, B., Fackrell, K., 2019. One size does not fit all: developing common standards for outcomes in early-phase clinical trials of sound-, psychology-, and pharmacology-based interventions for chronic subjective tinnitus in adults. Trends Hear. 23. https://doi.org/10.1177/2331216518824827.

Hara, K., Miura, I., Suzuki, M., Tanaka, T., 2024. Assessing future potentiality of technologies from the perspective of "imaginary future generations" – a case study of hydrothermal technology. Technol. Forecast. Soc. Chang. 202, 123289. https://doi.org/10.1016/j.techfore.2024.123289.

He, Y., Guo, S., Dong, P., Wang, C., Huang, J., Zhou, J., 2022. Techno-economic comparison of different hybrid energy storage systems for off-grid renewable energy applications based on a novel probabilistic reliability index. Appl. Energy 328, 120225. https://doi.org/10.1016/j.apenergy.2022.120225.

Hecht, E., Ehrhart, B., Groth, K., 2021. Hydrogen Plus Other Alternative Fuels Risk Assessment Models (HyRAM+) Version 4.0 Technical Reference Manual. https://doi.org/10.2172/1832082.

Hideg, É., Mihók, B., Gáspár, J., Schmidt, P., Márton, A., Báldi, A., 2021. Assessment in horizon scanning by various stakeholder groups using Osgood's semantic differential scale – a methodological development. Futures 126, 102677. https://doi.org/10.1016/j.futures.2020.102677.

IEA, 2020. Energy Technology Perspectives 2020. Special Report on Clean Energy Innovation. Accelerating Technology Progress for a Sustainable Future.

IEA, 2022. Research Gaps Towards Hydrogen Commercialization.

IEA, 2023a. World Energy Outlook 2023.

IEA, 2023b. Global Hydrogen Review 2023.

Jahel, C., Bourgeois, R., Bourgoin, J., Daré, W., De Lattre-Gasquet, M., Delay, E., Dumas, P., Le Page, C., Piraux, M., Prudhomme, R., 2023. The future of social-ecological systems at the crossroads of quantitative and qualitative methods. Technol. Forecast. Soc. Change 193. https://doi.org/10.1016/j.techfore.2023.122624.

Kashanizadeh, B., Mohammadnezhad Shourkaei, H., Fotuhi-Firuzabad, M., 2022. A bi-stage multi-objectives optimization of smart multi energy system with active consumers. Sustain. Prod. Consum. 31, 707–722. https://doi.org/10.1016/j.spc.2022.03.028.

Király, G., Köves, A., Pataki, G., Kiss, G., 2016. Assessing the participatory potential of systems mapping. Syst. Res. Behav. Sci. 33, 496–514. https://doi.org/10.1002/SRES.2374.

Köhler, J., Wietschel, M., Whitmarsh, L., Keles, D., Schade, W., 2010. Infrastructure investment for a transition to hydrogen automobiles. Technol. Forecast. Soc. Change 77, 1237–1248. https://doi.org/10.1016/j.techfore.2010.03.010.

Korteling, Johan.E. (Hans), Paradies, G.L., Sassen-van Meer, J.P., 2023. Cognitive bias and how to improve sustainable decision making. Front. Psychol. 14. https://doi.org/10.3389/fpsyg.2023.1129835.

Krumm, A., Süsser, D., Blechinger, P., 2022. Modelling social aspects of the energy transition: what is the current representation of social factors in energy models? Energy 239, 121706. https://doi.org/10.1016/J.ENERGY.2021.121706.

Kurtz, J., Sprik, S., Bradley, T.H., 2019. Review of transportation hydrogen infrastructure performance and reliability. Int. J. Hydrog. Energy 44, 12010–12023. https://doi.org/10.1016/j.ijhydene.2019.03.027.

Kurtz, J., Sprik, S., Peters, M., Bradley, T.H., 2020. Retail hydrogen station reliability status and advances. Reliab. Eng. Syst. Saf., 106823 https://doi.org/10.1016/j.ress.2020.106823.

Kwon, D., Choi, S.K., Yu, C., 2022. Improved safety by crossanalyzing quantitative risk assessment of hydrogen refueling stations. Int. J. Hydrog. Energy 47, 10788–10798. https://doi.org/10.1016/j.ijhydene.2021.12.211.

Leimeister, M., Kolios, A., 2018. A review of reliability-based methods for risk analysis and their application in the offshore wind industry. Renew. Sust. Energ. Rev. 91, 1065–1076. https://doi.org/10.1016/J.RSER.2018.04.004.

Li, W., Zhang, Leiqi, Lv, H., Zhang, Lijun, Liu, M., Zhang, C., He, P., 2023. Reliability analysis of hydrogen storage composite pressure vessel with two types of random-interval uncertainties. Int. J. Hydrog. Energy 48, 31685–31699. https://doi.org/10.1016/j.ijhydene.2023.04.312.

Liang, H., Pirouzi, S., 2024. Energy management system based on economic flexi-reliable operation for the smart distribution network including integrated energy system of hydrogen storage and renewable sources. Energy 293, 130745. https://doi.org/10.1016/j.energy.2024.130745.

Liu, T., Yang, Z., Duan, Y., Hu, S., 2022. Techno-economic assessment of hydrogen integrated into electrical/thermal energy storage in PV+ Wind system devoting to high reliability. Energy Convers. Manag. 268, 116067. https://doi.org/10.1016/j.enconman.2022.116067.

McCarthy, R., 2004. A Methodology to Assess the Reliability of Hydrogen-based Transportation Energy Systems.

McCarthy, R.W., Ogden, J.M., Sperling, D., 2007. Assessing reliability in energy supply systems. Energy Policy 35, 2151–2162. https://doi.org/10.1016/j.enpol.2006.06.016.

McDowall, W., 2012. Technology roadmaps for transition management: the case of hydrogen energy. Technol. Forecast. Soc. Change 79, 530–542. https://doi.org/10.1016/j.techfore.2011.10.002.

Moradi, R., Groth, K.M., 2019. Hydrogen storage and delivery: review of the state of the art technologies and risk and reliability analysis. Int. J. Hydrog. Energy 44, 12254–12269. https://doi.org/10.1016/j.ijhydene.2019.03.041.

Moustapha Mai, T., Cristofari, C., Azzaro-Pantel, C., 2023. Design and operation of hydrogen supply chains: a review on technology integration and system optimization. In: Hydrogen Economy: Processes, Supply Chain, Life Cycle Analysis and Energy Transition for Sustainability, 115–164. https://doi.org/10.1016/B978-0-323-99514-6.00002-9.

Müller, K., 2022. Effect of the reliability on the success of hydrogen technologies. Chem. Eng. Technol. 45, 385–388. https://doi.org/10.1002/CEAT.202100033.

Ohlendorf, N., Löhr, M., Markard, J., 2023. Actors in multi-sector transitions - discourse analysis on hydrogen in Germany. Environ. Innov. Soc. Trans. 47, 100692. https://doi.org/10.1016/j.eist.2023.100692.

Park, C., Lim, S., Shin, J., Lee, C.Y., 2022. How much hydrogen should be supplied in the transportation market? Focusing on hydrogen fuel cell vehicle demand in South Korea: hydrogen demand and fuel cell vehicles in South Korea. Technol. Forecast. Soc. Change 181, 121750. https://doi.org/10.1016/J.TECHFORE.2022.121750.

Park, J., Kang, S., Kim, S., Kim, H., Cho, H.-S., Lee, J.H., 2024a. Enhancing the economic viability and reliability of renewables based electricity supply through power-to-gas-to-power with green hydrogen. Energy Convers. Manag. 310, 118485. https://doi.org/10.1016/j.enconman.2024.118485.

Park, J., Kang, S., Kim, S., Kim, H., Kim, S.-K., Lee, J.H., 2024b. Optimizing green hydrogen systems: balancing economic viability and reliability in the face of supply-demand volatility. Appl. Energy 368, 123492. https://doi.org/10.1016/j.apenergy.2024.123492.

Raford, N., 2015. Online foresight platforms: evidence for their impact on scenario planning & strategic foresight. Technol. Forecast. Soc. Change 97, 65–76. https://doi.org/10.1016/J.TECHFORE.2014.03.008.

Sacio-Szymańska, A., Fantoni, G., Daheim, C., 2016. Foresight as a key enabler of innovation in the economy. Introduction to the topical collection. Eur. J. Futures Res. 4, 1–4. https://doi.org/10.1007/S40309-016-0104-2/METRICS.

Scholten, D., 2013. The reliability of energy infrastructures. In: The Organizational Requirements of Technical Operations, 14, pp. 173–205. https://doi.org/10.1177/178359171301400205 doi:10.1177/178359171301400205.

Seker, S., Aydin, N., 2022. Assessment of hydrogen production methods via integrated MCDM approach under uncertainty. Int. J. Hydrog. Energy 47 (5), 3171–3184. https://doi.org/10.1016/j.ijhydene.2021.07.232.

Sharma, G.D., Verma, M., Taheri, B., Chopra, R., Parihar, J.S., 2023. Socio-economic aspects of hydrogen energy: an integrative review. Technol. Forecast. Soc. Change 192, 122574. https://doi.org/10.1016/J.TECHFORE.2023.122574.

Sovacool, B.K., Axsen, J., Sorrell, S., 2018. Promoting novelty, rigor, and style in energy social science: towards codes of practice for appropriate methods and research design. Energy Res. Soc. Sci. 45, 12–42. https://doi.org/10.1016/j.erss.2018.07.007.

Su, H., Zio, E., Zhang, J., Li, X., Chi, L., Fan, L., Zhang, Z., 2019. A method for the multi-objective optimization of the operation of natural gas pipeline networks considering supply reliability and operation efficiency. Comput. Chem. Eng. 131, 106584. https://doi.org/10.1016/J.COMPCHEMENG.2019.106584.

Swiss Re Institute, 2022. De-risking the Hydrogen Economy.

Thakkar, N., Paliwal, P., 2024. Data driven MCDM models for reliability-economic-environmental analysis of energy storage based autonomous micro-grid. J. Energy Storage 81, 110408. https://doi.org/10.1016/j.est.2023.110408.

Tugnoli, A., Landucci, G., Cozzani, V., 2009. Key performance indicators for inherent safety: application to the hydrogen supply chain. Process. Saf. Prog. 28, 156–170. https://doi.org/10.1002/PRS.10303.

Turkcu, D., Tura, N., 2023. The dark side of sustainable packaging: battling with sustainability tensions. Sustain. Prod. Consum. 40, 412–421. https://doi.org/10.1016/J.SPC.2023.07.007.

van Bree, B., Verbong, G.P.J., Kramer, G.J., 2010. A multi-level perspective on the introduction of hydrogen and battery-electric vehicles. Technol. Forecast. Soc. Change 77, 529–540. https://doi.org/10.1016/j.techfore.2009.12.005.

van der Duin, P., Trott, P., Ortt, R., 2024. Failed technology management: introducing 'future technology myopia' and how to address it. Technol. Forecast. Soc. Change 198, 122927. https://doi.org/10.1016/j.techfore.2023.122927.

van de Kerkhof, M., Cuppen, E., Hisschemöller, M., 2009. The repertory grid to unfold conflicting positions: the case of a stakeholder dialogue on prospects for hydrogen. Technol. Forecast. Soc. Change 76, 422–432. https://doi.org/10.1016/j.techfore.2008.07.004.

Wang, J., Qi, M., Liu, S., Zhao, D., 2024. Comprehensive safety assessment of a hydrogen liquefaction system based on an integrated system-theoretic process analysis (STPA) and best-worst method (BWM). Int. J. Hydrog. Energy 66, 479–489. https://doi.org/10.1016/j.ijhydene.2024.04.082.

West, M., Al-Douri, A., Hartmann, K., Buttner, W., Groth, K.M., 2022. Critical review and analysis of hydrogen safety data collection tools. Int. J. Hydrog. Energy 47, 17845–17858. https://doi.org/10.1016/j.ijhydene.2022.03.244.

Wu, T., Wang, J., 2023. Reliability evaluation for integrated electricity-gas systems considering hydrogen. IEEE Trans. Sustain. Energy 14, 920–934. https://doi.org/10.1109/TSTE.2022.3229896.

Wu, T., Wang, J., Yue, M., 2022. On the integration of hydrogen into integrated energy systems: modeling, optimal operation, and reliability assessment. IEEE Open Access J. Power Energy 9, 451–464. https://doi.org/10.1109/OAJPE.2022.3204216.

Yodo, N., Afrin, T., Yadav, O.P., Wu, D., Huang, Y., 2023. Condition-based monitoring as a robust strategy towards sustainable and resilient multi-energy infrastructure systems. Sustain. Resilient Infrastruct. 8, 170–189. https://doi.org/10.1080/23789689.2022.2134648.

Yue, M., Lambert, H., Pahon, E., Roche, R., Jemei, S., Hissel, D., 2021. Hydrogen energy systems: a critical review of technologies, applications, trends and challenges. Renew. Sust. Energ. Rev. 146, 111180. https://doi.org/10.1016/j.rser.2021.111180.

Sofía De-León Almaraz is an Associate Professor at Corvinus University of Budapest, Hungary, where she has been part of the Department of Supply Chain Management since 2021. Her focus has been on hydrogen-related topics since 2011, and she holds a PhD in Process and Environmental Engineering from the *Institut National Polytechnique de Toulouse*. She is an active member of the first European Hydrogen Sustainability and Circularity Panel of the Clean Hydrogen Partnership. Her current research interests include hydrogen, sociotechnology, reliability, life cycle and sustainability assessment, and futures studies.

Tchougoune Moustapha Mai is a research engineer at the University of Corsica, working in the Renewable Energies Department of the *Laboratoire Sciences pour l'Environnement* (SPE). His PhD research area is based on the optimization of hydrogen supply chain deployment in island territories. The author has a master's in engineering in energy systems from the University of the French West Indies. He collaborated for 4 years with the *Laboratoire de Genie Chimique* in Toulouse, France on the optimization of hydrogen supply chain.

Iris Rocio Melendez is a consultant on risk management with >15 years' experience in industrial and financial fields (e.g., Hedge Funds, Stock Exchange, and Leasing Global Financing). She holds a degree in Industrial Engineering and a master's in risk management from EALDE (Spain). Experienced in leading ISO 9001 implementation, and development of quality control and IT security strategies. Facilitator of ERP implementation. Her contributions extend to academia, where she has taught Operations Research at the Industrial Engineering School in Costa Rica. Currently she works as a consultant on risk management at Corvinus University of Budapest.

M.K. Loganathan is a Professor at the Suresh Gyan Vihar University, India, researching in Reliability, EV Batteries, Sustainable Transportation, and Supply Chain Management. He was associated with The Assam Kaziranga University, Assam, as Dean (SET), Head of Mechanical Engineering and Director (IQAC) and currently as a Research Faculty. He obtained PhD (Reliability Engg) from IIT (Indian Institute of Technology Delhi). He undertook postdoctoral studies at Chang Gung University, Taiwan. He worked for a leading auto industry in India in Reliability Engg. Currently, he is acting as a Chief Reliability Consultant at the Crimson Energy Expert Pvt. Ltd., New Delhi.

Catherine Azzaro-Pantel is a Professor at *Laboratoire de Genie Chimique* from the *Institut National Polytechnique de Toulouse* (INPT-ENSIACET), in charge of a cross-cutting axis on energy. Co-founder of a Master-level program in "EcoEnergy" at INPT-ENSIACET based on an original and equilibrated combination of *process systems engineering* and *electrical engineering* disciplines, with an interdisciplinary problem-solving approach necessary for identifying sustainable solutions in the energy sector. Her research interests lie in *Process Systems Engineering* with specific focus on optimization methods for process design, i.e., chemical engineering, applied mathematics, operations research, and optimization. Author of several articles and books on hydrogen systems.