

A Hybrid Reliability Assessment Framework for Green Hydrogen Production Systems under Data Scarcity

Leonardo Andrade Santos

Centro de Tecnologia e Geociências, Universidade Federal de Pernambuco, Brazil. E-mail: leonardo.las@ufpe.br

Caio Bezerra Souto Maior

Centro de Tecnologia e Geociências, Universidade Federal de Pernambuco, Brazil. E-mail: caio.maior@ufpe.br

The rapid expansion of green hydrogen production has intensified the need for reliable and resilient system architectures capable of operating under highly uncertain conditions. Despite increasing attention to economic and environmental aspects, the operational reliability of green hydrogen plants remains insufficiently addressed in the literature, largely due to the scarcity of historical failure data. This work proposes a hybrid reliability assessment framework that integrates qualitative knowledge from the literature with stochastic modeling techniques to evaluate the operational reliability of solar-based green hydrogen production systems. The methodology combines Failure Mode and Effects Analysis (FMEA) for systematic identification and classification of failure mechanisms, Fault Tree Analysis (FTA) to model causal relationships leading to system unavailability, and Monte Carlo simulation to propagate uncertainty throughout the system. In the absence of deterministic failure data, failure probabilities are modeled using Weibull distributions, whose parameters are inferred from qualitative FMEA inputs and physical analogies, while uncertainty is explicitly represented through stochastic sampling. The framework is applied to a representative photovoltaic-powered PEM electrolyzer system, focusing on the top event defined as unplanned shutdown or capacity reduction. The proposed approach demonstrates how dispersed qualitative information can be transformed into quantitative reliability indicators, even under severe data scarcity. The framework provides actionable insights for asset management, maintenance prioritization, and early-stage design decisions, supporting the development of more robust and reliable green hydrogen production plants.

Keywords: Green hydrogen, Reliability analysis, Fault Tree Analysis, Failure Mode and Effects Analysis, Monte Carlo simulation, Data scarcity, Renewable energy, Photovoltaic–electrolyzer integration, System availability, Risk Analysis, Stochastic modelling.

1. Introduction

The increasing international pressure to decarbonize productive sectors and to expand the share of renewable sources in the global energy matrix has intensified the search for safe, efficient, and sustainable solutions for energy storage and transportation (Chen et al. 2025). In this context, green hydrogen (GH₂) has emerged as a promising alternative, both for hard-to-abate industrial sectors and for the storage of surplus energy generated from intermittent renewable sources. (Marzouk 2024)

Despite the growing discussion regarding its economic feasibility and environmental benefits, the operational reliability of green hydrogen production systems remains a relatively underexplored topic in the technical and academic literature. Furthermore, the scarcity of historical

failure data from green hydrogen plants significantly limits the direct application of classical quantitative reliability methods (Tuhi et al. 2024)

On the other hand, a large body of qualitative information on photovoltaic systems and electrolyzers is dispersed throughout the technical literature and remains underutilized (Santos et al., 2024). This work addresses this gap by proposing a hybrid methodology that integrates Failure Mode and Effects Analysis (FMEA) for knowledge structuring, Fault Tree Analysis (FTA) for logical failure modeling, and Monte Carlo Simulation (MCS) for stochastic quantification. The proposed framework transforms qualitative literature-based information into quantitative decision-support metrics, providing a prioritized risk map to

support asset management, mitigate operational risks, and enhance the availability of future green hydrogen production plants.

2. Methodology

This section presents the methodological framework adopted to evaluate the reliability of green hydrogen production systems. The proposed approach integrates qualitative and quantitative techniques in a sequential manner, allowing the analysis of system behavior even under conditions of limited operational data.

To conduct the reliability analysis, this work focuses on a representative GH₂ production architecture coupled with a photovoltaic (PV) power generation system. For the purpose of generalizing the analysis and concentrating on the most critical elements, the system was modeled considering its essential subsystems only, as illustrated in Figure 1

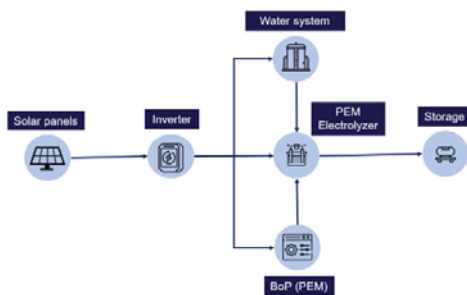


Figure 1: Critical elements system

2.1. FMEA structuring

After defining the components that constitute the system under study, a comprehensive literature review was conducted. This review aimed to systematically map the most recurrent failure modes reported for each system component.

In this FMEA, each failure mode was classified according to its causes and effects, as well as its impact on personnel safety, severity, and occurrence. In addition, the Criticality Index ($Cr = S \times O$) was adopted as an auxiliary metric for risk ranking, enabling the identification of failure modes with higher mitigation priority under real

operating conditions. During the literature review, no consistent or sufficiently detailed information was found to support the inclusion of the detectability criterion typically used in conventional FMEA formulations. Consequently, the analysis was based on the Criticality Index, which is widely adopted in scenarios characterized by data scarcity and limited operational feedback.

2.2. FTA structuring

The quantification of system reliability was performed using the FTA methodology, with logical structures derived directly from the failure modes identified in the FMEA. The top event analyzed in the fault tree was defined as “Unexpected Shutdown or Reduction in Hydrogen Production”, with emphasis on operational reliability and asset availability.

Given the uncertainty inherent to the input parameters—derived from literature sources rather than precise field data—a traditional analytical approach based solely on mean occurrence values would be insufficient and potentially misleading, as it would conceal the variability associated with system risk. Therefore, uncertainty propagation through the fault tree was carried out using MCS.

In total, the FMEA identified 24 failure modes, which were translated into 26 basic events and 13 intermediate events within the fault tree structure.

2.3. Fault tree quantification method

The central objective of this methodology is to transition from a qualitative assessment to a quantitative stochastic modeling approach. While the FMEA provides the identification of failure modes, their causes, effects, and severity levels, the FTA enables the understanding of the logical interactions among these events and the propagation of their probabilities up to system collapse (Maior et al. 2022; Macedo et al. 2023).

The methodological innovation proposed in this work lies in replacing point estimates of failure with probability distributions. By modeling each basic event through a statistical distribution, the analysis explicitly incorporates the theoretical and physical assumptions embedded in its parameters.

The justification for the universal adoption of the Weibull distribution to model all failure modes is based on its adaptability and physical consistency. Widely recognized in the literature as a versatile distribution, the Weibull distribution is capable of representing the three phases of the bathtub curve—infant mortality, useful life, and wear-out—solely through the adjustment of the shape parameter β . In this way, it mathematically unifies distinct failure behaviors, reducing to an exponential distribution when $\beta = 1$ (random failures) and approaching a normal-like behavior when $\beta > 2$ (wear-out mechanisms), without the need to alternate between different statistical families.

This property allows physical degradation mechanisms observed in real systems, such as corrosion or electronic component failure, to be directly translated into the temporal domain of the simulation. Accordingly, the information extracted from the FMEA—particularly occurrence levels and failure types—is used to estimate the physical and temporal parameters of the distributions, enabling the calculation of failure probabilities for each basic event.

The analysis horizon was defined as a five-year mission time, a critical period that encompasses the plant’s consolidation phase and reflects a realistic scenario in which long-term historical data are nonexistent or scarce. The quantification process was structured into three sequential stages and executed through a MCS, as detailed in the following subsections.

2.3.1. Failure Physics Modeling (β Parameter)

The first step of the quantification process consists of defining the physical behavior of failures over time. Based on the detailed FMEA analysis, it was possible to classify the underlying stress and degradation mechanisms associated with each failure mode.

The Weibull shape parameter (β) is the mathematical indicator that represents this failure physics, determining whether the failure rate is decreasing, constant, or increasing over time. In this study, it was assumed that early-life failures related to manufacturing defects or human errors during installation—typical of the infant mortality region ($\beta < 1$)—are effectively mitigated through

rigorous quality control and commissioning procedures. Therefore, the proposed model focuses exclusively on the useful life and wear-out phases of the bathtub curve.

To assign values to the β parameter in the absence of destructive testing data, a physical analogy approach was adopted. Each failure mode was classified into three categories based on classical reliability literature (Abernethy, 2006; O’Connor; KLEYNER, 2012)

Recognizing that the exact failure physics of each component may vary according to operating conditions, β was not treated as a fixed point estimate. Instead, it was modeled as a random variable described by a Beta distribution, whose parameters were elicited using a PERT-based approach. Table 1 summarizes the adopted parameters for each physical failure category identified in the FMEA.

Table 1. Physical Failure Categories and Weibull Shape Parameter (β) Definition

Physical Category	Failure Behavior	PERT Parameters for β (Min – Mode – Max)
Electronics / External Events	Constant Failure Rate (Random)	0.9 – 1.0 – 1.1
Physicochemical Degradation	Increasing Failure Rate	2.0 – 3.0 – 3.5
Thermomechanical Fatigue	Increasing Failure Rate	2.5 – 3.25 – 4.0

Source: Developed by the author (2025).

2.3.2. Temporal Parameterization (α Parameter)

Once the physical nature of the failure has been defined, the second step consists of sizing the temporal magnitude of the risk, that is, determining the scale parameter (α) of the Weibull distribution.

The main challenge lies in translating the ordinal Occurrence (O) rating from the FMEA into a physical temporal quantity. The fundamental assumption adopted in this study is that the FMEA risk classification reflects the cumulative probability of failure P_{life} at the end of the component’s nominal service life t_{nom} , as

specified in technical manuals and reliability literature.

Given the inherent subjectivity of classifications such as “High” or “Low” occurrence, and the absence of historical operational data for the plant, the probability P_{life} was not associated with point estimates. Instead, it was treated as a random variable under maximum uncertainty.

To this end, a standard Beta distribution, defined over the [0,1] domain, was adopted and parameterized using a PERT-based approach:

- **Bounds (Minimum and Maximum):**
Fixed at 0% and 100% for all occurrence levels. This choice avoids probabilistic dogmatism. By allowing the distribution to span nearly the entire probability spectrum (0 to 1), the model acknowledges the possibility—however remote—that a component classified as “Rare” may experience a premature catastrophic failure, or that a “Critical” component may exceed its expected service life. This ensures that the simulation explores tail scenarios that would otherwise be ignored by truncated distributions.
- **Central Tendency:**
The distinction among risk levels (1 to 5) is established exclusively through the mode of the distribution, shifting the probability mass according to the severity of the classification (see Table 2).

Table 2. α parametrization

FMEA rating (O)	Classification	PERT Parameters (Min – Mode – Max)
1	Remote	0.00 – 0.10 – 1.00
2	Low	0.00 – 0.30 – 1.00
3	Moderate	0.00 – 0.50 – 1.00
4	High	0.00 – 0.70 – 1.00
5	Critical	0.00 – 0.90 – 1.00

Source: Developed by the author (2025).

Given t_{nom} , the sampled probability P_{life} , and the central value of the shape parameter β , the scale parameter α becomes the only unknown in the Weibull cumulative distribution function (CDF). For each Monte Carlo iteration, the algorithm solves the following relationship:

$$P_{life} = 1 - e^{-\left(\frac{t_{nom}}{\alpha}\right)^\beta} \tag{1}$$

Solving for α :

$$\alpha = \frac{t_{nom}}{[-\ln(1 - P_{life})]^{1/\beta_{central}}} \tag{2}$$

In this manner, the qualitative uncertainty embedded in the FMEA is systematically converted into uncertainty regarding the actual durability of each component.

2.3.3. Stochastic Integration

The third and final stage of the methodology constitutes the computational core of the study, where the synthesis between failure physics (Stage 1) and temporal estimation (Stage 2) is performed. The objective is to compute the effective failure probability of each event over the critical analysis horizon $t_{target} = 5 \text{ years}$ and to propagate this uncertainty through the system logic.

Unlike deterministic approaches based on single-point estimates, this stage operates on simulation vectors. For each of the 100,000 Monte Carlo iterations, the algorithm executes the following procedures:

2.3.3.1. Parametric Recombination (Failure Probability Evaluation)

At this stage, the final coupling of uncertainties is performed. A deliberate decoupling strategy is adopted to ensure numerical stability and physical interpretability:

- In Stage 2, only the modal value of the Weibull shape parameter ($\beta_{central}$) was used to anchor the estimation of the scale parameter α , preventing shape variability from distorting the temporal calibration.
- In Stage 3, the full stochastic representation of the shape parameter β , generated in Stage 1, is reintroduced into the model.

As a result, for each Monte Carlo iteration i , the failure probability calculation incorporates:

- α_i , which embeds uncertainty associated with failure frequency derived from the FMEA occurrence rating;
- β_i , which captures uncertainty related to the physical aggressiveness of the degradation mechanism.

The failure probability of each basic event at the target time horizon is then computed using the Weibull CDF:

$$P_i(t_{target}) = 1 - e^{-\left(\frac{t_{target}}{\alpha_i}\right)^{\beta_i}} \quad (3)$$

2.3.3.2. Logical Propagation (Bottom-Up Aggregation)

Once the probability has been generated for all basic events, system-level propagation toward the top event is performed. This aggregation strictly follows the Boolean logic defined in the FTA applying the exact probabilistic formulations associated with each logical gate (AND, OR).

The bottom-up propagation is executed within each Monte Carlo iteration, ensuring that uncertainty is consistently transmitted from component level to subsystem and system levels.

2.3.3.3. Model Outputs and Uncertainty Diagnostics

The iterative execution of this process produces probability distributions for all basic, intermediate, and top events. This approach provides a deeper diagnostic capability, allowing:

- identification of dispersion and asymmetry in system-level failure probabilities;
- assessment of uncertainty amplification across the fault tree structure;
- support for risk-informed decision-making based on probabilistic ranges rather than deterministic indicators.

By embedding uncertainty at both the component and system levels, the proposed framework delivers a comprehensive probabilistic characterization of operational risk, even under severe data scarcity conditions.

3. Results and Discussion

3.1. FMEA Results

3.3.1. Qualitative Analysis

The qualitative FMEA reveals strong systemic coupling between the photovoltaic (PV) subsystem and the electrolyzer, indicating that failures in electricity generation propagate beyond local effects and directly affect hydrogen production reliability. Within the PV subsystem, several failure modes act as failure propagators rather than isolated events. Encapsulant

delamination, for example, facilitate the ingress of moisture and contaminants, increasing the likelihood of secondary electrical and corrosion-related failures. The analysis of junction boxes and mounting structures reveals a risk profile characterized by low occurrence but very high severity. Although these failures are infrequent ($O = 1-2$), their consequences include fire and electric shock hazards ($S = 4-5$). This combination results in high intrinsic risk, currently masked by low failure frequency.

The inverter was identified as the most critical component within the PV subsystem, presenting the highest risk priority index due to the combination of maximum severity ($S = 5$) and high occurrence ($O = 4$). Acting as the central conversion unit, inverter failure results in complete and immediate loss of photovoltaic generation, as consistently reported in the literature. For the green hydrogen plant, this leads to sudden loss of self-generated power, forcing system shutdown or full reliance on the grid, increasing operating costs and exacerbating electrolyzer cycling.

The FMEA of the PEM electrolyzer reveals a highly interconnected and critical risk profile. All major stack components—membrane, catalyst, and bipolar plates—exhibited high criticality indices. Notably, operational intermittency was the only failure mode assigned the maximum occurrence rating ($O = 5$), reflecting its dual origin: unavoidable variability from the solar cycle and controllable but difficult-to-manage operational start–stop events. This persistent intermittency induces dynamic stress conditions, including fluctuations in pressure, voltage, and current, accelerating degradation of stack components.

Failures in electrolyzer auxiliary systems—such as cooling, power conditioning and control, hydrogen purification and drying, gas separation, and water supply—significantly impact operational efficiency and safety. Cooling and water supply failures present high criticality due to risks of membrane overheating, irreversible

stack damage, and explosive scenarios. Power control system failures also show high criticality, as they hinder responses to abnormal conditions, leading to electrical and pressure-related hazards. Failures in hydrogen purification and drying mainly affect hydrogen quality and increase explosion risks.

3.3.2 Quantitative Analysis

The quantitative analysis indicates that 50% of the evaluated failure modes were classified as critical, highlighting the high relevance of mechanisms associated with operational safety and system integrity. All failures related to the electrolyzer were considered critical, whereas only 16.67% (two failures) were associated with the photovoltaic subsystem. Overall, 67% of the failure modes were rated as having high or critical criticality, confirming the significant impact of these failures on the system's overall performance.

Among the failure modes classified as medium risk, a predominance of the photovoltaic subsystem was observed, accounting for seven of the eight identified cases.

The FMEA analysis from a human safety perspective reveals an equally critical risk scenario. More than 53% of the identified failure modes present a high or critical impact on operator integrity. Subsystem-level stratification demonstrates a parity of critical risk between the photovoltaic and electrolysis subsystems, each presenting four critical failure modes and a comparable number of high-risk modes. This finding reinforces the need for an integrated safety management approach with equally rigorous protocols for both electrical hazards and flammable process safety.

3.2. FTA Results

The probabilistic Fault Tree Analysis focused on the top event defined as unplanned shutdown or capacity reduction of the plant, representing a loss of operational continuity rather than a catastrophic safety event. Simulation results indicate a high likelihood of operational disruption within the first five years, with a median cumulative probability (P50) of 87.41%. Even under optimistic conditions (P10 = 79.84%), the probability of interruption remains elevated, while pessimistic scenarios reach 92.89% (P90). The absence of a meaningful low-

probability tail (see Figure 2) indicates that uninterrupted operation without additional maintenance or redundancy is statistically unlikely, with full success probabilities ranging only from approximately 7% to 21%.

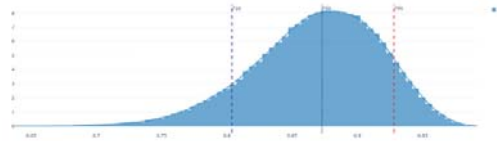


Figure 2: Top Event Histogram

The system reliability profile shows a rapid decline, reaching 50% by the end of the second year, indicating that the first significant failure requiring intervention is expected around this time. The uncertainty band remains relatively stable (15–20%) throughout the period, demonstrating that input uncertainty does not obscure the underlying degradation trend. Overall, the results reveal a combination of residual early-life failures and accelerated wear-out, refuting the assumption that the plant can operate under a low-maintenance regime during its initial five years.

4. Conclusion

This work developed a hybrid reliability assessment framework for GH2 production systems under conditions of severe data scarcity. By integrating FMEA, FTA, and MCS, the methodology transformed dispersed qualitative knowledge from the literature into quantitative reliability indicators suitable for decision-making. The results demonstrate that the system exhibits a structurally high risk of unavailability, with a median probability of interruption or capacity reduction of approximately 87% over a five-year horizon.

Overall, the results indicate that the operational viability of green hydrogen plants depends less on incremental technological improvements in electrolyzers and more on the robustness of the system architecture. Transitioning from demonstration projects to industrial-scale deployment requires the systematic introduction of redundancy in critical subsystems and energy storage solutions to mitigate intermittency.

5. Acknowledgement

The authors thank the Brazilian research funding agencies Human Resources Program (PRH 38.1) entitled “Risk Analysis and Environmental Modeling in the Exploration, Development and Production of Oil and Gas”, financed by ‘Agência Nacional de Petróleo (ANP)’ and managed by FAPESP, the ‘Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq): 402761/2023-5’, FACEPE: APQ-0504-3.08/25 and ‘Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)’ - Finance Code 001 for the financial support through research grants.

6. References

- Abernethy, Robert B. 2006. *The New Weibull Handbook: Reliability & Statistical Analysis for Predicting Life, Safety, Risk, Support Costs, Failures, and Forecasting Warranty Claims, Substantiation and Accelerated Testing, Using Weibull, Log Normal, Crow-AMSA, Probit, and Kaplan-Meier Models*. 5. ed. R.B. Abernethy.
- Chen, Jie, Xin Chen, Jie Sun, Jingkuo Qu, Xiangjiu Guan, e Shaohua Shen. 2025. “Thermal Management Matters in Photovoltaic–Electrocatalysis for Solar Hydrogen Production”. *EcoEnergy* 3 (2): 205–16. <https://doi.org/10.1002/ece2.84>.
- Macedo, July Bias, Caio Souto Maior, Isis Didier Lins, et al. 2023. “Using Experts’ Opinion for Bayesian Prior Reliability Distribution of on-Demand Equipment: A Case Study of a Novel Sliding Sleeve Valve for Open-Hole Wells”. *Reliability Engineering & System Safety* 238 (outubro): 109430. <https://doi.org/10.1016/j.ress.2023.109430>.
- Maior, Caio B.S., July B. Macêdo, Isis D. Lins, et al. 2022. “Bayesian Prior Distribution Based on Generic Data and Experts’ Opinion: A Case Study in the O&G Industry”. *Journal of Petroleum Science and Engineering* 210 (março): 109891. <https://doi.org/10.1016/j.petrol.2021.109891>.
- Marzouk, Osama A. 2024. “Expectations for the Role of Hydrogen and Its Derivatives in Different Sectors through Analysis of the Four Energy Scenarios: IEA-STEPS, IEA-NZE, IRENA-PES, and IRENA-1.5°C”. *Energies* 17 (3): 646. <https://doi.org/10.3390/en17030646>.
- O’Connor, Patrick D. T., and Andre Kleyner. 2012. *Practical Reliability Engineering*. 5th ed. Chichester: John Wiley & Sons.
- Santos, Leonardo Andrade, Caio Souto Maior, e Isis Didier Lins. s. d. *Revisão de Técnicas de Simulação Aplicadas na Produção de Hidrogênio Verde*. <https://doi.org/doi.org/10.71190/2024-12-1224691>.
- Tuhi, Farhana Yasmine, Marta Bucelli, e Yiliu Liu. 2024. “Technical failures in green hydrogen production and reliability engineering responses: Insights from database analysis and a literature review”. *International Journal of Hydrogen Energy* 94 (dezembro): 608–25. <https://doi.org/10.1016/j.ijhydene.2024.11.129>.