GREEN ECONOMICS
Vol.3, No.2, 2025, pp.65-95
https://doi.org/10.62476/ge.32.65

AI-DRIVEN RENEWABLE ENERGY ATLAS AND FORECASTING PLATFORM FOR GREEN HYDROGEN PRODUCTION IN ALGERIA

Kada Ahmed Yasser[1], Mehdaoui Abdelkhaleq[1], Guezouli Lyamine[1], Bahloul Derradji[1,2], Ouchen Rima[1]

[1]LEREESI, Laboratory, HNS-RE2SD, Batna, Algeria
[2]University of Sciences and Technology Houari Boumediene, Algiers, Algeria

Abstract. Green hydrogen is increasingly recognized as a cornerstone for decarbonizing energy systems, yet its large scale deployment requires accurate, regionally adaptive forecasting and robust economic evaluation. Leveraging open-access NASA POWER meteorological data and advanced machine learning, a novel platform is developed to forecast and simulate green hydrogen production potential across Algeria's diverse climatic zones. The system integrates Light Gradient Boosting Machine (LightGBM) models for solar irradiance forecasting demonstrating superior accuracy over deep learning alternatives and a dynamic electrolyzer simulation environment. Four renewable energy sourcing scenarios (100% solar, 100% wind and two hybrid mixes) are evaluated across seven representative regions, revealing pronounced spatial disparities in hydrogen yields. Southern provinces such as Adrar and Tamanrasset exhibit the highest solar-driven production, while hybrid configurations enhance resilience in semi-arid and northern areas. The study further incorporates a techno-economic analysis, providing Levelized Cost of Hydrogen (LCOH) estimates and policy insights aligned with Algeria's national hydrogen roadmap. Discussion extends to environmental and social implications of large-scale hydrogen deployment. The proposed platform offers a scalable, data-driven decision-support tool for site selection, infrastructure planning and export-oriented policy formulation, advancing Algeria's ambition to become a regional leader in green hydrogen.

Keywords: *Green hydrogen, AI forecasting, techno-economic analysis, LightGBM, electrolyzer simulation, NASA POWER, renewable energy policy, Algeria.*

Corresponding Author: *Guezouli Lyamine, LEREESI, Laboratory, HNS-RE2SD, Batna, Algeria, e-mail: lyamine.guezouli@hns-re2sd.dz*

Received: *4 April 2025;* ***Accepted:*** *22 May 2025;* ***Published:*** *30 May 2025.*

1. Introduction

As the global energy transition accelerates, green hydrogen produced via electrolysis powered by renewable sources is gaining momentum as a strategic energy carrier. Its ability to decarbonize hard to electrify sectors such as heavy industry, long-haul transport and seasonal storage makes it a cornerstone of future low-carbon systems. Yet, its viability depends not only on renewable potential but also on the ability to predict, optimize and scale production dynamically. Algeria boasts one of the most substantial renewable energy potentials in the Mediterranean region, primarily due to its vast solar and wind resources. The southern provinces, particularly in the Sahara, experience exceptional solar irradiance, with annual sunshine durations reaching up to 3,900 hours (Wikipedia contributors, 2024). This high solar exposure translates to significant energy yields, making regions like Tamanrasset, Adrar and In Salah ideal for photovoltaic and solar-thermal applications. In addition to solar resources, Algeria's High Plateaus and coastal zones offer complementary wind energy potential. Studies have identified areas such as Tiaret and El Bayadh as having favorable wind regimes, with average wind speeds at 10 meters elevation ranging between 5.6 and 6.3 m/s (Bouchair *et al*., 2023). These wind conditions support the development of hybrid renewable systems and reinforce Algeria's diversified energy portfolio. Recognizing these advantages, the Algerian government introduced its National Hydrogen Strategy in March 2023. The strategy sets a phased trajectory aiming to establish Algeria as a key player in the global green hydrogen market by 2040. It outlines pilot projects from 2023-2030, followed by production scaling and large-scale exports targeting 30-40 TWh/year by 2040 an estimated 10% of Europe's hydrogen needs (Institute for Sustainable Alternatives, 2023). Despite this vision, many studies and pilot programs have remained concentrated in a few southern wilayas, notably Adrar and Tamanrasset, neglecting the country's broader climatic and infrastructural heterogeneity. Not all regions share the same irradiance, wind profiles, water availability or proximity to energy infrastructure. A narrow geographic focus limits the accuracy and generalizability of national-scale planning. To address this, a spatially comprehensive, AI-integrated platform is needed one capable of simulating hydrogen yields across diverse locations under varying environmental and technical conditions. This article proposes such a system and demonstrates its application across seven representative Algerian regions.

The core problem lies in operationalizing this national potential under real world constraints. Solar and wind availability are highly variable and without accurate forecasting and adaptive system planning, hydrogen yields may remain inconsistent. Additionally, infrastructure limitations, water availability and policy uncertainty further complicate deployment decisions. To support meaningful national-scale planning, Algeria requires a digital solution that can simulate hydrogen production across multiple regions using realistic, time varying inputs.

This study proposes such a platform: an AI-powered forecasting and simulation system that leverages open-source NASA POWER data and advanced machine learning models to estimate hydrogen yields. The model is applied to seven representative regions across Algeria spanning the Sahara, High Plateaus and coastal zones and evaluates hydrogen production under four energy sourcing scenarios: 100% solar, 100% wind and two hybrid mixes.

In the following sections, we first review the state-of-the-art in AI-driven forecasting, electrolyzer optimization, catalyst innovation and regional green hydrogen planning, situating our work within the broader scholarly landscape. We then detail the architecture and methodological framework of the proposed forecasting and simulation platform, including data sources, machine learning models and system integration. Subsequent sections present and analyze simulation results across seven representative Algerian regions, highlighting spatial disparities and scenario-based hydrogen yields. We further assess the economic feasibility and policy implications of large-scale green hydrogen deployment, emphasizing cost drivers, infrastructure needs and strategic alignment with Algeria's national roadmap. Finally, we discuss the broader environmental and social impacts and outline future research directions to enhance the platform's predictive capabilities and policy relevance.

2. Related Work

The integration of artificial intelligence (AI) into green hydrogen systems has attracted growing scholarly and industrial interest, spanning diverse areas such as energy forecasting, electrolyzer optimization, catalyst design and regional policy planning. This section categorizes and critically reviews existing literature along four major thematic axes to contextualize the contributions of this study.

2.1. AI-Based Forecasting and Modeling for Green Hydrogen Production

In the field of Deep Learning for Wind and Solar Energy Forecasting in Hydrogen Production, recent advances in deep learning have significantly enhanced the forecasting of renewable energy outputs, a critical component for green hydrogen systems given the inherent variability of solar and wind resources. For instance, Alkhayat et al. explore how AI techniques can effectively model complex, nonlinear patterns in energy data to improve forecast accuracy (Nikulins *et al*., 2024). Building on this, Alhussan et al. (2023) propose a hybrid ensemble approach that combines BER and PSO algorithms to optimize RNN hyperparameters, leading to improved solar hydrogen production forecasts. Other researchers have introduced more spatially aware models, such as the hybrid GCN-LSTM framework by Liao et al. (2021) which captures both spatial and temporal dependencies in renewable energy datasets, thereby improving short term prediction accuracy for production planning. Similarly, Bosma and Nazari (2021) apply computer vision techniques using a ResNet-based CNN to extract spatio temporal features from weather maps, offering a novel angle for predicting solar and wind power outputs. Meanwhile, Li et al. (2022) advance the discussion on data privacy in distributed systems by proposing a federated deep reinforcement learning model that enables decentralized wind power forecasting without compromising data confidentiality.

Despite these advancements, many of these approaches remain focused on isolated prediction tasks and do not integrate forecasting with hydrogen production simulations or region-specific operational contexts. Most notably, they fall short of demonstrating real-time deployment using open-access satellite data or extending their utility into system-level frameworks that couple prediction with electrolyzer performance across diverse geographies. This persistent gap is precisely what our study addresses.

Table 1. AI-Based Forecasting and Modeling for Green Hydrogen Production

Focus Area	Representative Approaches	Key Gaps Identified
Forecasting solar and wind availability for hydrogen planning	LSTM, GCN-LSTM hybrids, BER-PSO-RNN ensembles, CNNs on weather maps, federated learning	Lack of integration with hydrogen yield simulation; real-time open-source data underutilized; limited regional validation

While forecasting plays a pivotal role in anticipating energy availability, it alone is insufficient for ensuring efficient hydrogen production. Once the availability of renewable energy is anticipated, intelligent scheduling and optimization of the electrolysis process become essential to maximize hydrogen yield and system efficiency. The following section examines how AI is leveraged to enhance the performance and reliability of hydrogen production systems.

2.2. AI-Driven Optimization of Electrolyzer Performance and Hydrogen Production Efficiency

Artificial intelligence (AI) plays a crucial role in optimizing electrolyzer performance and hydrogen production efficiency across multiple fronts. For instance, Veettil Muhammed (2023) illustrate how AI techniques can enhance operational control by managing variable inputs to improve system reliability and efficiency, though such implementations remain largely confined to controlled environments and lack real-time regional adaptation. Similarly, Alhussan et al. (2023) emphasize the value of AI in predictive maintenance for electrolyzers, enabling early failure detection and reduced downtime. However, while effective at the component level, these methods are not integrated with broader forecasting or production optimization systems. On the materials side, K E and Padhan (2024) apply machine learning to accelerate electrocatalyst discovery by predicting catalytic performance, a contribution that, while important, remains peripheral to system-level production modeling. Lesniak et al. (2024) advance the field further by developing AI-driven dynamic scheduling strategies that enhance grid flexibility and improve energy utilization in hydrogen systems, yet these models often assume deterministic energy supply, overlooking real-time variability in renewable resources. Kumar and Singh (2023) propose an integrated AI-based framework for process level optimization of green hydrogen production, aiming to maximize efficiency and reduce cost; however, such frameworks typically rely on idealized datasets or commercial-scale assumptions, limiting their use in open-access or regional simulation contexts. Collectively, these contributions demonstrate AI's broad potential in hydrogen system optimization. Yet, few of them link predictive modeling with real time, region specific simulations powered by publicly accessible meteorological datasets. Our study aims to bridge this gap by embedding AI into an open-access, Algeria-specific simulation platform for hydrogen production planning.

Table 2. AI-Driven Optimization of Electrolyzer Performance and Hydrogen Production Efficiency

Focus Area	Representative Approaches	Key Gaps Identified
Optimizing hydrogen production system performance	AI-based predictive maintenance, dynamic electrolyzer scheduling, system-level optimization frame-works	Assumptions of deterministic input; lack of real-time data integration; absence of open-access regional platforms

Beyond system level optimization, recent advances in AI have also targeted material-level challenges particularly in the discovery and tuning of catalysts that govern electrochemical efficiency. Catalysts play a fundamental role in determining the efficiency of hydrogen production, particularly under intermittent energy input conditions. The next section explores how AI is transforming catalyst discovery and material optimization for hydrogen systems.

2.3. AI in Catalyst Development and Materials Science

AI is increasingly applied in the discovery and optimization of catalysts, which are essential for efficient hydrogen production. Researchers at the University of Toronto, for instance, developed a novel catalyst recipe using AI to improve hydrogen fuel production efficiency (Abed *et al*., 2024). This showcases the potential of intelligent systems in catalyst innovation, though the real-world scalability of such discoveries remains underexplored. Similarly, the integration of AI with quantum computing has been investigated to accelerate binary photocatalyst discovery for hydrogen production (Wayo *et al*., 2024), yet these approaches are still in their early stages and are constrained by computational intensity and limited experimental validation. Machine learning has also proven useful for predicting electrocatalytic performance in hydrogen evolution reactions (K E & Padhan, 2024), but the pathway from prediction to experimental proof and industrial application continues to pose significant challenges. Optimization strategies driven by AI have focused on improving catalyst stability and activity under varying operational conditions (Yin *et al*., 2025); however, questions persist regarding the compatibility of these catalysts with existing production systems. Data-driven models for catalyst design have furthered the identification of high performing materials (Lai *et al*., 2024), although issues such as training data limitations, model transparency and generalizability across diverse catalysts remain. Despite this progress, these AI-driven developments are largely confined to material science domains and lack direct coupling with operational hydrogen production frameworks. Bridging this gap requires interdisciplinary collaboration to translate catalyst-level innovations into integrated, system-scale deployments for green hydrogen infrastructure.

Table 3. AI in Catalyst Development and Materials Science

Focus Area	Representative Approaches	Key Gaps Identified
Catalyst discovery and performance enhancement	Machine learning for electrocatalyst prediction, AI-quantum computing integration, data-driven catalyst optimization	Mostly confined to lab-scale; weak coupling with hydrogen systems; limited industrial validation

While the technological frontiers of forecasting, optimization and catalyst design are advancing globally, their deployment must be grounded in regional realities. In Algeria, a set of national initiatives and policy frameworks are beginning to shape the country's hydrogen trajectory.

2.4. Regional Studies and Policy Initiatives in Algeria

Focusing on Algeria, several studies and national initiatives have assessed the country's potential for green hydrogen production and the integration of AI in energy planning. A simulation-based analysis using HOMER Pro compared photovoltaic productivity and hydrogen yields across Algerian regions, revealing strong potential in both desert and non-desert zones (Benchenina *et al*., 2025). While technically promising, this study underscores the need to tackle water resource limitations and cost barriers to enable scalability. Complementing these findings, Algeria's National Renewable Energy and Hydrogen Plan aims to reduce dependency on fossil based electricity and expand the role of renewables (Energy Box, 2025). Although the plan outlines steps to improve institutional and technical capacities, its effectiveness will depend on sustained policy commitment and cross-border cooperation. Strategic efforts like the SouthH2 Corridor further demonstrate Algeria's ambition to export clean hydrogen to Europe particularly Germany and Italy through transcontinental pipelines (Hydrogen Central, 2025). Still, this initiative requires significant infrastructure upgrades and financial mobilization to reach operational readiness. Algeria's National Hydrogen Strategy, introduced in 2023, sets a long-term vision to position the country as a global hydrogen hub by 2040 (International Solar Alliance, 2023), but the phased roadmap it proposes must be backed by consistent execution and investment. Finally, a field study in the Ouargla region evaluated hydrogen generation from PV systems under desert conditions, offering empirical validation for system viability (Gougui *et al*., 2019), yet also pointing to environmental constraints and logistical complexities in scaling such operations. Collectively, these contributions affirm Algeria's strategic position in the green hydrogen landscape, while highlighting that success will depend on overcoming infrastructural deficits, strengthening implementation capacity and nurturing durable international alliances.

Table 4. Regional Studies and Policy Initiatives in Algeria

Focus Area	Representative Approaches	Key Gaps Identified
National green hydrogen planning and policy	HOMER Pro simulations, SouthH2 Corridor, national hydrogen strategy, pilot studies in desert regions	Insufficient infrastructure, policy uncertainty and weak adoption of AI - based planning tools

Together, these four categories provide a comprehensive overview of the technical, material and policy dimensions of AI-driven green hydrogen systems. However, a persistent gap remains in integrating these layers into a unified, region-specific platform. Our study addresses this gap by developing a unified, region-specific AI-enabled platform tailored to Algeria's green hydrogen development priorities.

3. Methodology

This work adopts a systems-oriented methodology that combines real-time environmental data acquisition, predictive modeling using AI and hydrogen output simulation. The overall process involves the integration of open-access climate datasets, machine learning-based forecasting for solar irradiance and a modular simulation of hydrogen production using a representative electrolyzer model. The approach comprises the following phases:

3.1. Simulation and Scenario Design

A simplified model of a Proton Exchange Membrane (PEM) electrolyzer is used to simulate daily hydrogen production. The model receives solar and wind energy inputs (in kWh/day) and converts them to hydrogen mass (kg/day) based on fixed conversion efficiency and an electrolyzer oversizing factor (OSF). Although the platform architecture supports multiple electrolyzer types and user-defined configurations, this study focuses on one predefined configuration to ensure comparability across modeled scenarios. The implications of this choice are discussed in the Limitations section.

We simulate four renewable energy input scenarios:

- 100% solar;
- 100% wind;
- 70% solar / 30% wind;
- 50% solar / 50% wind.

Due to performance instability in the wind forecasting models (detailed in Section 6.4), wind energy data used in these simulations is derived directly from historical satellite observations without AI-based forecasting.

3.2. Forecasting Using AI Models

The forecasting module in the platform currently focuses on solar irradiance prediction using a LightGBM model. This component enables anticipatory simulation of green hydrogen output based on forecasted solar availability.

For solar forecasting, the LightGBM model is trained on historical satellite-derived data from NASA POWER. The feature set includes:

- Lagged values of Global Horizontal Irradiance (GHI);
- Daily mean temperature;
- Day-of-year (encoded);
- Location-specific geographic coordinates.

The output is a one-day-ahead prediction of solar irradiance, which is then used to estimate green hydrogen production under different regional and energy configuration scenarios.

In parallel, we have explored predictive modeling of wind resources - specifically wind speed (WS50M) and wind direction (WD50M) - using deep learning architectures including LSTM, GRU and Transformer. Although these models were tested and performance metrics computed (see Section 8.5), the results remain unsatisfactory due to high temporal variability and inconsistency in wind data. Consequently, wind forecasting is not yet integrated into the production simulations and is left for future system updates. At present, wind input for hydrogen modeling is derived from historical (non-forecasted) data.

3.3. Platform Implementation

The platform is fully developed in Python. Data ingestion and preprocessing pipelines use pandas and $pvlib$. Forecasting is implemented using $LightGBM$ and models are serialized via $joblib$. The simulation backend is exposed via a REST API built with Flask. The user interface allows input specification, forecasting execution, scenario simulation and spatial visualization of hydrogen output across Algeria.

3.4. Evaluation and Scenario Comparison

To assess both forecasting accuracy and hydrogen production potential, we perform region-specific simulations across four representative Algerian cities: Algiers, Skikda, Ouargla and Tamanrasset. These regions represent coastal, highland and desert climate zones. The solar forecasting model is evaluated using standard metrics such as Mean Absolute Error (MAE) and coefficient of determination (R^2). Hydrogen outputs under each energy scenario are then compared across regions to analyze spatial and configuration-dependent variation.

4. Green hydrogen simulation using real-time data

Green hydrogen is produced through the electrolysis of water, a process that uses electricity from renewable sources to split water (H_2O) into hydrogen (H_2) and oxygen (O_2). In this study, we simulate this process using real-time meteorological data - specifically solar irradiance and wind speed - sourced from NASA's POWER database, combined with a dynamic electrolyzer model tailored to Algerian environmental conditions.

4.1. Green Hydrogen Production Process

Green hydrogen is produced through a clean and sustainable process that relies on renewable energy sources. The key steps are as follows:

1. Electricity Generation from Renewable Sources:

- **Solar Power:** Photovoltaic (PV) panels convert sunlight into direct current (DC) electricity. This electricity can be used immediately or stored in batteries.
- **Wind Power:** Wind turbines convert kinetic energy from wind into alternating current (AC) electricity, typically fed into the grid or storage systems.
- **Hybrid System (Solar + Wind):** In hybrid setups, both solar and wind sources are integrated. A smart energy management system balances their output based on availability (e.g., solar during the day, wind at night). This improves reliability and reduces fluctuations in power supply to the electrolyzer.

2. Electrolysis: The electricity generated is supplied to an electrolyzer, which splits water (H_2O) into hydrogen (H_2) and oxygen (O_2) using electrochemical reactions:

$$2H_2O(l) \longrightarrow 2H_2(g) + O_2(g)$$

Modern electrolyzers include Proton Exchange Membrane (PEM), Alkaline or Solid Oxide technologies, each suited for different energy input conditions.

3. Hydrogen Collection and Compression: The produced hydrogen gas is collected, filtered and typically compressed for storage or transport. Compression increases energy density and facilitates its use in industrial or mobility applications.

4. Storage and Distribution: Hydrogen can be stored in pressurized tanks, underground caverns or converted to carriers like ammonia. It is then distributed via pipelines, trailers or fueling stations.

5. End-Use Applications: Green hydrogen is used in various sectors such as clean transportation (fuel cells), industrial heating, chemical production (e.g., ammonia) or grid energy storage.

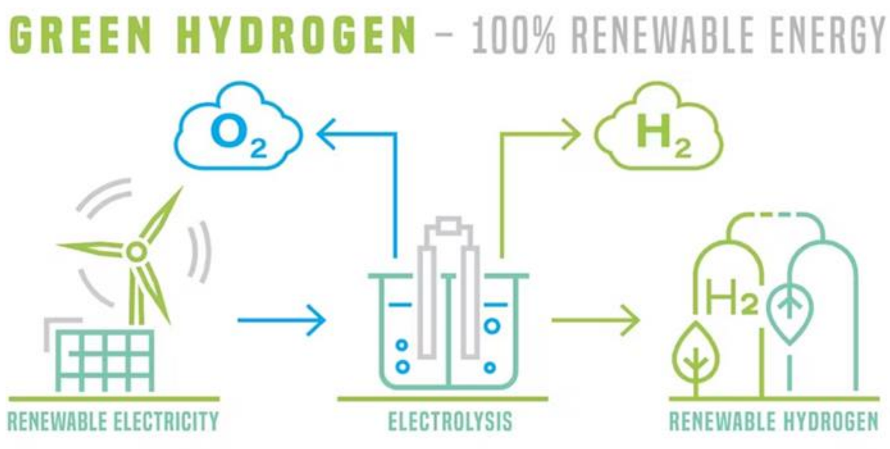

Figure 1. Green Hydrogen production steps (13)

4.2. Real-Time Data Integration

We retrieved hourly solar irradiance (Global Horizontal Irradiance, GHI) and wind speed data for multiple Algerian regions using the NASA POWER API. These datasets represent the real-time availability of renewable energy resources and are used as direct inputs for green hydrogen potential estimation.

The data acquisition process includes:

- **Solar inputs:** GHI in W/m^2.
- **Wind inputs:** Wind speed at 50m height (m/s), wind Direction at 50m, scaled for turbine height.
- **Time resolution:** Hourly for daily simulations, with scenarios built for seasonal variation.

4.3. Electrolyzer Simulation Framework

A Proton Exchange Membrane (PEM) electrolyzer model is implemented to simulate hydrogen production. PEM technology is selected due to its rapid response to fluctuating power inputs and compatibility with intermittent renewables. The electrolyzer converts electrical energy E_{elec} (from solar or wind) into hydrogen mass m_{H_2}, governed by the following formula:

$$m_{H_2} = \frac{\eta \cdot E_{elec}}{HHV_{H_2}}$$

Where:

- m_{H_2}: Hydrogen mass produced (kg);
- η: Electrolyzer efficiency (dimensionless, typically 0.6-0.7);
- E_{elec}: Electrical energy input (kWh);
- HHV_{H_2}: Higher heating value of hydrogen (≈ 39.4 kWh/kg).

This formula is used for each hour of the simulation, accumulating the hourly hydrogen production over the day.

4.4. Production Scenarios

Four simulation scenarios were evaluated to compare the effect of different energy source combinations on hydrogen yield:

- **Scenario A - 100% Solar:** Hourly GHI is converted to PV energy using a standard module efficiency of 18%.
- **Scenario B - 100% Wind:** Hourly wind data converted to power output via a wind turbine power curve.
- **Scenario C - 70% Solar / 30% Wind Hybrid:** Weighted combination of PV and wind energy inputs.
- **Scenario D - 50% Solar / 50% Wind Hybrid:** Balanced hybrid configuration.

Each scenario simulates hourly hydrogen output based on real meteorological data and electrolyzer response, allowing for a comparative assessment of performance across Algerian regions.

5. Mathematical Formulation of Green Hydrogen Production

The simulation logic is built to estimate the **annual hydrogen production** ($H_{2,annual}$) in kilograms or tonnes, based on site-specific renewable energy potential and electrolyzer parameters. The process is divided into distinct modules for solar and wind energy, which are then aggregated according to the chosen energy scenario.

5.1. Solar Energy Output Estimation

POA Irradiance Calculation:

$$G_{POA} = G_{b,POA} + G_{d,POA} + G_{g,POA}$$

Where:
G_{POA} = Total Plane of Array irradiance [W/m^2];
$G_{b,POA}$ = Beam (direct) irradiance on the plane of array [W/m^2];
$G_{d,POA}$= Diffuse irradiance on the plane of array [W/m^2];
$G_{g,POA}$= Ground-reflected irradiance on the plane of array [W/m^2].

Cell Temperature (Faiman Model):

$$T_{cell} = T_{amb} + \frac{G_{POA}}{G_{NOCT}}(T_{cell,NOCT} - T_{amb,NOCT})(1 - \frac{\tau\alpha}{\eta_{el}})$$

Where:
T_{cell} = Cell temperature [°C];
T_{amb} = Ambient temperature [°C];
G_{NOCT} = Reference irradiance at NOCT, typically 800 W/m^2;
$T_{cell,NOCT}$ = Nominal Operating Cell Temperature [°C];
$T_{amb,NOCT}$ = Ambient temperature at NOCT [°C];
$\tau\alpha$ = Effective transmittance-absorptance product (unitless);
η_{el} = Electrical efficiency of the PV module (unitless).

Hourly PV Output:

$$P_{PV,hourly} = G_{POA,hourly} \cdot A_{panel} \cdot \eta_{panel} \cdot (1 - L_{temp}) \cdot L_{other}$$

Where:
$P_{PV,hourly}$ = Hourly PV power output [W];
$G_{POA,hourly}$ = Hourly plane-of-array irradiance [W/m^2];
A_{panel} = Total area of PV panels [m^2];
η_{panel} = Efficiency of the PV module (unitless);
L_{temp} = Loss factor due to temperature (unitless);
L_{other} = Other system losses (wiring, inverter, etc.) (unitless).

Annual Energy per MW:

$$E_{PV,annual\ per\ MW} = \sum_{t=1}^{8760} P_{PV,hourly} \cdot PR \quad \text{[MWh/MW/year]}$$

Where:
$E_{PV,annual\ per\ MW}$ = Annual energy output per MW of PV installed [MWh];
PR = Performance ratio accounting for all losses (unitless).

5.2. Wind Energy Output Estimation

Wind Speed Extrapolation to Hub Height:

$$\nu_{hub} = \nu_{ref} \cdot \left(\frac{h_{hub}}{h_{ref}}\right)^{\alpha}$$

Where:
ν_{hub} = Wind speed at hub height [m/s];
ν_{ref} = Wind speed at reference height [m/s];
h_{hub} = Hub height of the wind turbine [m];
h_{ref} = Reference measurement height [m];
α = Wind shear exponent (unitless), typical value 0.14-0.3.

Annual Wind Energy Output:

$$E_{Wind,annual\ per\ MW} = CF_{wind} \cdot P_{rated} \cdot 8760$$

Where:
$E_{Wind,annual\ per\ MW}$ = Annual energy output per MW installed [MWh];
CF_{wind} = Capacity factor of the wind turbine (unitless);
P_{rated} = Rated power of wind turbine [MW] 8760 = Total number of hours in a year.

5.3. Total Renewable Energy Input

100% PV:

$$E_{RE,total} = C_{PV} \cdot E_{PV,annual\ per\ MW}$$

Where:
$E_{RE,total}$ = Total annual renewable energy [MWh];

C_{PV} = Installed PV capacity [MW].

100% wind:

$$E_{RE,total} = C_{Wind} \cdot E_{Wind,annual\ per\ MW}$$

Where:
C_{Wind} = Installed wind capacity [MW].

Hybrid System:

$$C_{RE} = OSF \cdot C_{Ely}, \quad C_{PV} = f_{PV} \cdot C_{RE}, \qquad C_{Wind} = f_{Wind} \cdot C_{RE}$$

Where:
OSF = Oversize factor for renewable energy system;
C_{Ely} = Capacity of the electrolyzer [MW];
f_{PV} , f_{Wind} = Fraction of total RE capacity allocated to PV or Wind.

$$E_{RE,total} = C_{PV} \cdot E_{PV,annual\ per\ MW} + C_{Wind} \cdot E_{Wind,annual\ per\ MW}$$

5.4. Electrolyzer Energy Use and Sizing

$$E_{Ely,annual} = E_{RE,total} \cdot \eta_{grid\rightarrow ely}, \qquad CF_{ely} = \frac{E_{Ely,annual}}{C_{Ely} \cdot 8760}$$

Where:
$E_{Ely,annual}$ = Annual energy input to electrolyzer [MWh];
$\eta_{grid\rightarrow ely}$ = Efficiency of energy conversion from grid to electrolyzer (unitless);
CF_{ely} = Capacity factor of the electrolyzer (unitless).

5.5. Hydrogen Output Estimation

$$H_{2,annual}[kg] = \frac{E_{Ely,annual} \cdot 1000}{SEC_{H_2}}, \qquad H_{2,annual}[tonnes] = \frac{H_{2,annual}[kg]}{1000}$$

Where:
$H_{2,annual}$ = Annual hydrogen production [kg or tonnes];
SEC_{H_2} = Specific energy consumption of electrolyzer [kWh/kg H_2];
1000 = Conversion factor from MWh to kWh or kg to tonnes.

6. Forecasting and prediction using AI

To support anticipatory management of green hydrogen systems, this study incorporates artificial intelligence (AI) models to forecast key environmental variables driving hydrogen production. The forecasting framework targets two primary sources: solar irradiance and wind resources (wind speed and direction). For solar forecasting, several models were tested - including Random Forest, LSTM, CNN and Transformer architectures with LightGBM ultimately selected for deployment due to its superior performance on historical satellite data from NASA POWER. For wind forecasting, experimental models based on deep learning (LSTM, GRU, Transformer) were developed and evaluated using 50-meter wind speed (WS50M) and wind direction (WD50M) data. While preliminary results are reported, significant variability in wind patterns and inconsistencies across data sources led to unstable model behavior, here is the full work with results:

6.1. Motivation for AI-Based Forecasting

Given the intermittent nature of renewable energy, particularly solar and wind, accurate forecasting is essential to ensure:

- Effective electrolyzer scheduling,
- Infrastructure planning for storage and transport,
- Avoidance of over-sizing or under-utilization of generation assets.

AI models provide a flexible, data-driven alternative to deterministic physical models, particularly in data-scarce environments.

6.2. Data Inputs and Feature Engineering

For each location, the following meteorological features are used as model inputs across the different forecasting models developed in this study:

- Global Horizontal Irradiance (GHI);
- Diffuse Horizontal Irradiance (DHI);

- Wind speed (mean, minimum, maximum);
- Wind direction;
- Air temperature and surface pressure;
- Solar elevation angle;
- Time-based features (day of year, month, sine/cosine encoding of date).

These variables are sourced from NASA POWER and other open-access datasets, ensuring spatial and temporal consistency across Algerian regions. Feature engineering includes lagged values to capture temporal dependencies and harmonic encoding of seasonal cycles. All inputs are normalized using Min-Max scaling prior to training.

6.3. SOLAR Preliminary Modeling Experiments:

Before settling on LightGBM as the core forecasting model, we conducted a series of experiments using three alternative AI approaches commonly employed in renewable energy prediction: Random Forest (RF), Long Short-Term Memory networks (LSTM) and Transformer based architectures. Each model was trained and evaluated on the same feature-engineered dataset using consistent splitting and validation strategies. The aim was to benchmark their respective performance, interpretability and suitability for integration into a real-time hydrogen forecasting platform.

6.3.1. Random Forest Model

Random Forest (RF) is a robust ensemble learning method that averages multiple decision trees to reduce variance and overfitting. In this work, the RF model was used as a baseline to forecast daily Global Horizontal Irradiance (GHI) based on engineered meteorological and temporal features. The model was trained on the same feature set and data split as subsequent models to ensure fair comparison.

- Mean Absolute Error (MAE): 0.437;
- Root Mean Squared Error (RMSE): 0.646;
- R-squared (R^2): 0.834.

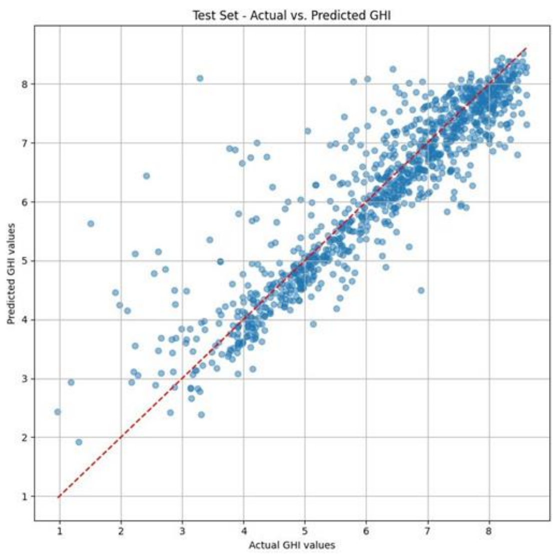

Figure 2. Random Forest Model – Actual vs. Predicted GHI on the Test Set

These results indicate a reasonably strong correlation between predicted and actual GHI values, although the RF model tends to underperform during high-variance or extreme weather conditions. Figure 2 shows the predicted vs. actual GHI values for the test set. Most points cluster around the ideal line, but the dispersion increases at higher GHI values, suggesting limitations in capturing outlier behavior.

6.3.2. Long Short-Term Memory (LSTM) Model

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) well-suited for capturing temporal dependencies in sequential data. In this experiment, an LSTM model was constructed to forecast GHI using a sliding window of historical weather and temporal features.

The model architecture consisted of stacked LSTM layers with dropout regularization to prevent overfitting. Training was conducted using the same input features and data splits as in the Random Forest model to ensure consistency.

- Mean Absolute Error (MAE): 0.596;
- Root Mean Squared Error (RMSE): 0.794;
- R-squared (R^2): 0.749.

While LSTM models are often powerful for time-series forecasting, they are computationally intensive and sensitive to hyperparameter tuning. Figure 3 displays the actual vs. predicted GHI values. The performance trend and deviations from the ideal diagonal line provide insights into the model's predictive accuracy across different irradiance levels.

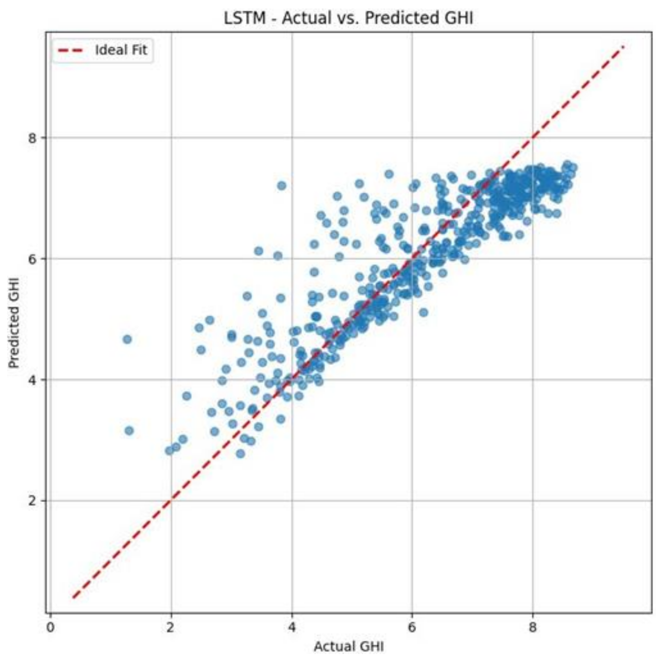

Figure 3. LSTM Model – Actual vs. Predicted GHI on the Test Set

6.3.3. Transformer-Based Model

Transformer architectures, initially developed for natural language processing, have recently been adapted to time series forecasting due to their ability to capture long-range dependencies via self-attention mechanisms. In this setup, a sequence-to-sequence

Transformer was trained to predict GHI based on multi-day historical inputs.

The Transformer used a multi-head attention encoder-decoder framework and was trained with the same feature set and target variables as the other models.

- Mean Absolute Error (MAE): 0.518;
- Root Mean Squared Error (RMSE): 0.769;
- R-squared (R^2): 0.765.

Despite its modeling power, the Transformer required significantly more training time and memory and its generalization capacity in limited-data settings was a point of concern. Further commentary will be added once performance trends are fully analyzed. Figure 4 shows predicted versus actual GHI values. Visual alignment with the identity line will be discussed once the results are available.

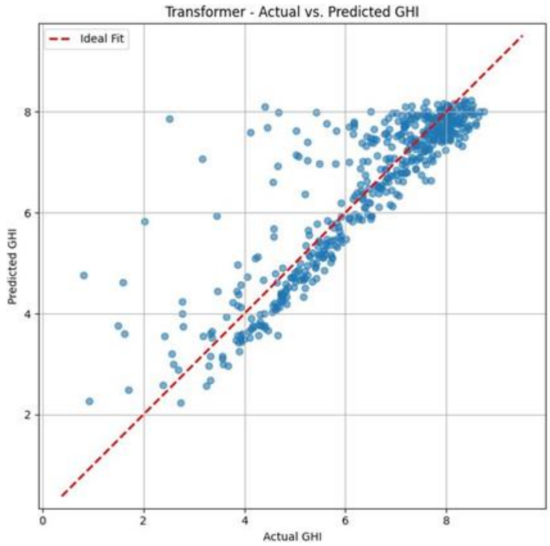

Figure 4. Transformer Model - Actual vs. Predicted GHI on the Test Set

6.3.4. Convolutional Neural Network (CNN) Model

Convolutional Neural Networks (CNNs), though originally designed for image processing, have been effectively adapted for time-series forecasting by treating sequential data as structured arrays. Their capacity to extract localized temporal features makes them suitable for predicting meteorological variables such as GHI (Borovykh *et al*., 2017). In this experiment, a 1D-CNN architecture was implemented, comprising multiple convolutional and pooling layers followed by dense layers. The model was trained on the same preprocessed time-series inputs as the other models, using lagged GHI and meteorological features. This architecture aimed to capture short-term temporal dependencies while maintaining computational efficiency.

- Mean Absolute Error (MAE): 0.611;
- Root Mean Squared Error (RMSE): 0.813;
- R-squared (R^2): 0.738.

The CNN model captured short-term temporal dependencies effectively but was less suited for long-term forecasting due to its localized receptive field. Figure 5 presents the model's predictive accuracy across GHI values.

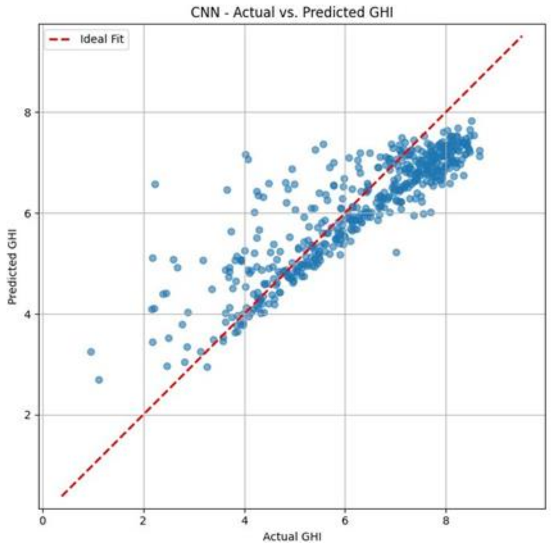

Figure 5. CNN Model – Actual vs. Predicted GHI on the Test Set

6.3.5. Final Evaluation of LightGBM on the Test Set

Among the tested models, LightGBM demonstrated the highest accuracy and generalization capability across all evaluation metrics, as shown in Table 5. It proved particularly effective in capturing daily and seasonal variability in solar irradiance, while maintaining low computational overhead.

Figure 6 illustrates the model's predictive behavior on the test set. The close alignment of predicted values with the identity line confirms the model's suitability for deployment in real-time hydrogen forecasting workflows.

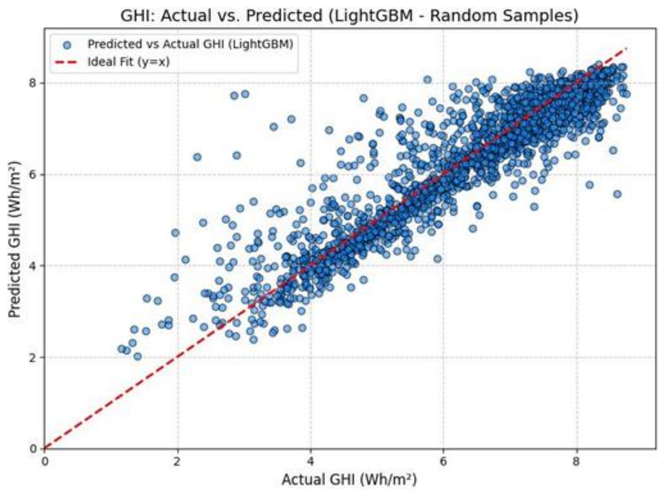

Figure 6. GHI: Actual vs. Predicted Using LightGBM Model (Random Samples)

6.4. WIND Preliminary Modeling Experiments

To assess the feasibility of AI-based wind forecasting, we evaluated three deep learning models - Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) and Transformer - using historical satellite data. The goal was to predict two wind parameters critical for hydrogen production: wind speed at 50 meters (WS50M) and wind direction at 50 meters (WD50M).

Each model was trained on the same lagged meteorological features and evaluated with consistent data splits. The performance metrics below reflect unscaled values for interpretability.

6.4.1. Transformer-Based Model

Transformer models demonstrated relatively better performance in capturing wind direction variability.

- **WD50M - MAE:** 30.604, **RMSE:** 63.699, $\mathbf{R^2}$**:** 0.625;
- **WS50M - MAE:** 1.343, **RMSE:** 1.688, $\mathbf{R^2}$**:** 0.261.

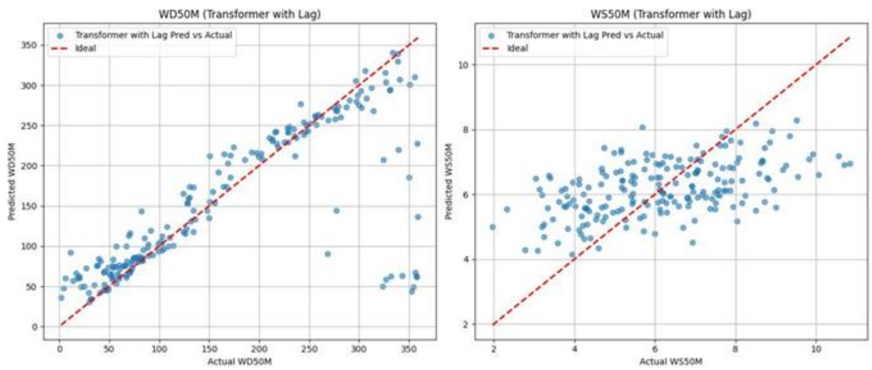

Figure 7. Transformer Model – Predicted vs. Actual for WS50M and WD50M

6.4.2. 4-Layer LSTM Model

LSTM showed slightly inferior results but with stable convergence during training.

- **WD50M - MAE:** 34.072, **RMSE:** 65.305, $\mathbf{R^2}$**:** 0.606;
- **WS50M - MAE:** 1.358, **RMSE:** 1.706, $\mathbf{R^2}$**:** 0.245.

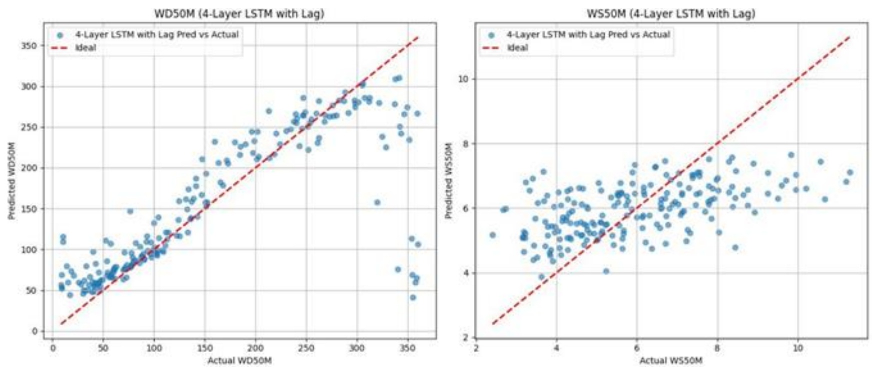

Figure 8. LSTM Model - Predicted vs. Actual for WS50M and WD50M

6.4.3. 4-Layer GRU Model

The GRU model produced similar results but exhibited more variance in wind direction predictions.

- **WD50M - MAE:** 33.431, **RMSE:** 66.411, $\mathbf{R^2}$**:** 0.593;
- **WS50M - MAE:** 1.356, **RMSE:** 1.704, $\mathbf{R^2}$**:** 0.247.

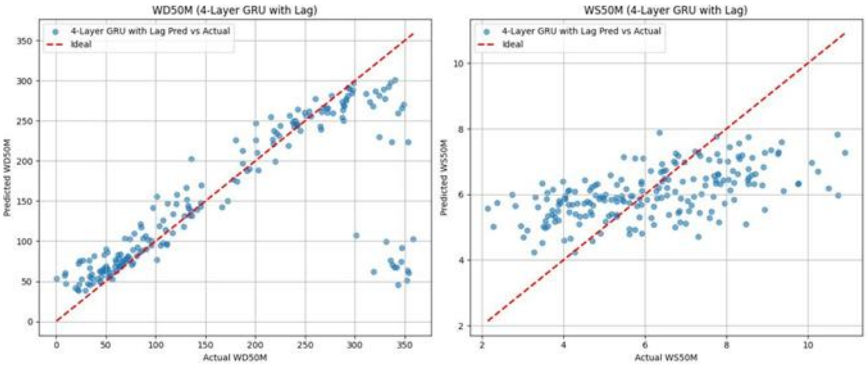

Figure 9. GRU Model – Predicted vs. Actual for WS50M and WD50M

Summary: Despite moderate performance in wind direction estimation, the models underperformed in wind speed prediction due to high variability and inconsistent patterns across regional datasets. As a result, wind forecasting is not yet integrated into the simulation pipeline. Future work will aim to refine data preprocessing and explore spatiotemporal ensemble models.

6.4.4. Model Evaluation and Comparison

To assess the forecasting capabilities of the different models, we compared their performance using three key metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Coefficient of Determination (R2). These metrics provide a comprehensive understanding of both average prediction accuracy and model robustness under varying conditions.

Solar Forecasting Models:

- **Random Forest** offered a stable baseline with moderate predictive strength and fast training time, though it lacked sensitivity to temporal patterns in the data.
- **LSTM**, as a recurrent neural network, showed improved temporal awareness but required careful tuning and was more sensitive to overfitting.
- **Transformer** architectures demonstrated promising modeling capacity for long-term dependencies, yet their training complexity and computational overhead limited their practical use in this context.
- **CNN** provided faster training convergence and handled short-term pattern recognition well, but it underperformed in capturing seasonal variability compared to tree-based models.
- **LightGBM** emerged as the most efficient and reliable model overall. It consistently achieved high accuracy while maintaining fast training speed and low memory usage. Its ability to handle nonlinearities and feature interactions made it well-suited for the structured meteorological data used in this study (Lara-Ben´ıtez *et al*., 2021).

Wind Forecasting Models:

- **LSTM, GRU and Transformer** models were tested to forecast both wind speed (WS50M) and wind direction (WD50M). Transformer slightly outperformed the recurrent networks in terms of MAE and R^2 for wind direction, though overall predictive power remained moderate.

- While the Transformer model achieved the best R^2 score of 0.625 for wind direction, performance on wind speed remained limited across all models (maximum R^2 of 0.261).
- Across the board, wind forecasting models faced challenges due to high variability, frequent fluctuations and inconsistency in regional wind data. These issues impacted training stability and generalization, especially in arid and coastal zones.
- As a result, wind forecasting models were not integrated into the simulation workflow at this stage, although their evaluation remains valuable for future enhancement of the platform.

6.4.5. Interpretation

Solar Forecasting: While all solar models demonstrated moderate-to-strong predictive performance, LightGBM outperformed others across all key metrics. It also required significantly less computational overhead than deep learning approaches, making it more practical for scalable deployment. Furthermore, LightGBM's ability to handle tabular data with missing or highly correlated features contributed to its robustness across different regions and seasons. Based on these findings, the LightGBM model was selected as the core forecasting engine for the solar prediction pipeline.

Wind Forecasting: Deep learning models such as Transformer, LSTM and GRU showed acceptable performance for wind direction prediction (WD50M), with the Transformer achieving an R^2 of 0.625. However, wind speed (WS50M) predictions remained less accurate due to inherent data noise and instability in time-series patterns. Given these limitations, wind forecasting was not incorporated into production simulations. The wind component currently relies on historical datasets, but future work will focus on improved spatiotemporal modeling and transfer learning techniques to enhance predictive robustness.

6.5. Model Structure for Solar Forcasting

The chosen model is a **LightGBM Regressor**, characterized by:

- Gradient boosting framework based on decision trees,
- Histogram-based algorithm for faster training and lower memory usage,
- Leaf-wise tree growth with depth constraints for better accuracy,
- Optimized using mean squared error (MSE) as the loss function.

Separate models were trained for predicting:

- Daily PV-based energy potential,
- Daily wind-based energy potential.

These predictions are then used in conjunction with the electrolyzer model to estimate forecasted hydrogen output.

6.6. Solar Forecasting Workflow

1. **Data Acquisition:** Retrieve historical and real-time weather data from NASA POWER API.
2. **Feature Construction:** Apply transformations and lags.
3. **Prediction:**

$$\hat{E}_{solar}(t), \hat{E}_{wind}(t) \xrightarrow{LGBM\ Models} \widehat{m}H_2(t)$$

4. **Post-Processing:** Rescale the forecasted energy values and compute hydrogen yield using the efficiency model.

6.7. Model Training and Validation

6.7.1. Data Splitting Strategy

To ensure robust performance evaluation, the dataset was chronologically split as follows:
- **Training Set:** 70% of the earliest data;
- **Validation Set:** 15% of the middle portion;
- **Test Set:** 15% of the most recent data.

Time-based splitting was chosen to mimic real-world forecasting scenarios and prevent data leakage.

All preprocessing steps were fit only on the training data and applied to the other sets.

6.7.2. Model Tuning and Hyperparameters

The LightGBM model was tuned using a randomized grid search with 5-fold cross-validation applied to the training and validation sets. The key hyperparameters were optimized as:
- $\mathrm{num\_leaves} = 31$;
- $\mathrm{learnining\_rate} = 0.05$;
- $\mathrm{n\_estimators} = 500$;
- $\mathrm{max\_depth} = -1$;
- $\mathrm{subsample} = 0.8$;
- $\mathrm{colsample\_bytree} = 0.8$;
- $\mathrm{reg\_alpha} = 0.1$ (L1 regularization);
- $\mathrm{reg\_lambda} = 0.2$ (L2 regularization).

Early stopping was applied with a patience of 30 rounds to prevent overfitting. The model’s best iteration was selected based on minimum validation RMSE.

6.7.3. Cross-Validation Results

The average RMSE over 5-fold cross-validation on the validation set was 0.65 Wh/m^2, confirming consistent performance. The final model was retrained on both the training and validation sets before evaluation on the test set.

6.7.4. Solar Model Robustness

The relatively high R^2 score of 0.84 on unseen test data suggests that the model generalizes well across different seasons and geographical regions. Residual analysis further confirms the absence of major prediction bias.

7. Solar Data Processing

The effectiveness of the forecasting and simulation framework hinges on accurate preprocessing of meteorological data and its integration into a set of predefined renewable energy scenarios for green hydrogen production. This section outlines the preprocessing pipeline and the scenario logic used in the platform.

7.1. Data Acquisition and Cleaning

Meteorological data were retrieved from NASA's POWER API with hourly resolution for a full year across multiple Algerian *wilayas*. The following parameters were collected:

- Global Horizontal Irradiance (GHI),
- Direct Normal Irradiance (DNI),
- Wind speed at 50m,
- wind direction at 50m,
- Surface temperature and pressure.

Processing steps include:

- Temporal resampling to daily values for forecasting,
- Missing value imputation using linear interpolation,
- Wind speed extrapolation to turbine hub height using the power law,
- Solar data transposition to the plane of array using pvlib.

7.2. Regional Coverage

Simulations were performed across all the country here are several representative regions across Algeria:

- **South (e.g., Adrar, Tamanrasset):** High solar potential,
- **North (e.g., Algiers, Skikda):** More moderate solar but viable wind conditions,
- **Highlands (e.g., Batna):** Mixed resource profiles.

Each scenario was applied uniformly across these locations to ensure comparative insight into regional disparities in green hydrogen potential.

8. Results and Discussion

This section presents the outcomes of green hydrogen production simulations across seven regions in Algeria under four renewable energy sourcing scenarios. The results combine deterministic simulations based on real-time NASA POWER data and AI-based solar energy forecasts generated using a LightGBM model. Wind energy predictions are currently in development and will be integrated in future work with best results.

8.1. Regional Hydrogen Production Scenarios and Insights

Figure 10 presents the annual hydrogen production potential across seven Algerian regions under four renewable energy configurations: 100% Solar, 100% Wind, 70% Solar / 30% Wind and 50% Solar / 50% Wind. These simulations are based on regional solar and wind resource profiles and assume an electrolyzer oversizing factor (OSF) of 2.0, which helps mitigate intermittency by allowing higher renewable capacity than nominal electrolyzer input. These results are specific to the selected electrolyzer type. The implications for other configurations supported by the platform are left for future work (see Limitations). The results reveal significant spatial variability:

– **Timimoun** and **Tamanrasset** achieve the highest hydrogen yields under the

100% solar scenario due to exceptional solar irradiance levels.

– **Djelfa** and **Batna** exhibit robust performance under both hybrid configurations, with 70% PV / 30% wind performing nearly as well as the solar-only case.

– **Alger** and **Tlemcen**, despite more moderate irradiance, benefit from hybrid setups, which improve reliability and seasonal resilience.

– Pure wind setups perform best in **Timimoun** and **Djelfa**, but fall short elsewhere, particularly in southern and coastal regions where wind resource is less consistent.

Hybrid strategies offer a valuable trade-off between total annual yield and energy reliability, especially in regions with complementary wind and solar profiles.

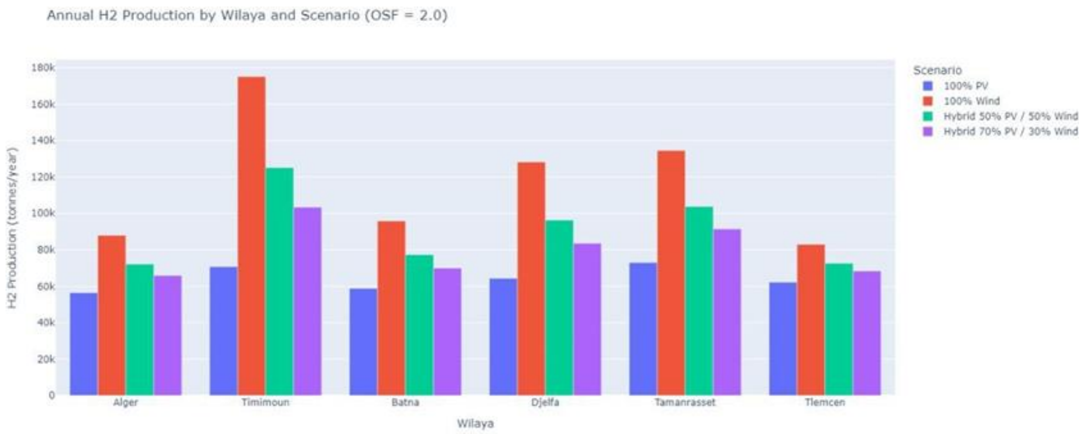

Figure 10. Annual Hydrogen Production (tonnes/year) Across Wilayas for Four Renewable Scenarios (OSF = 2.0)

8.2. Sustainability Indicators

Beyond energy output, sustainability metrics offer additional insights for evaluating the real-world viability of green hydrogen production. Based on literature benchmarks and electrolyzer specifications:

- **Water Use:** Assuming PEM electrolyzers consume approximately 9 to 10 liters of deionized water per kg of H (International Energy Agency, 2022), high-output regions like Adrar and Tamanrasset may require over 6,000 m^3/year for large-scale plants.
- **Land Use Efficiency:** With solar PV requiring roughly 1.5 to 2 hectares per MW of installed capacity (International Energy Agency, 2021) and yield estimates above 600 tons H/year in the south, solar-based hydrogen in desert zones achieves high production per hectare ratios.
- **CO Displacement Potential:** Replacing fossil-derived hydrogen (e.g., SMR) with forecasted green hydrogen yields could reduce CO emissions by 10 kg per kg H (International Energy Agency, 2023), translating to 6,000 to 8,000 tons of avoided emissions per site annually in top-performing zones.

8.3. AI-Based Solar Prediction Model Performance

To enable reliable and location-specific forecasting of solar energy availability, we developed and bench- marked several AI models for predicting Global Horizontal Irradiance (GHI), a key input for estimating solar energy and subsequently green hydrogen yield. The models evaluated included Random Forest (RF), Long Short-Term Memory networks (LSTM), Transformer-based architectures, Convolutional Neural

Networks (CNN) and LightGBM regressors.

Each model was trained on the same NASA POWER-derived meteorological features, using consistent preprocessing, feature engineering and temporal splitting strategies. Their performance was assessed based on the test set using three key metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Coefficient of Determination (R^2). Table 5 summarizes the results.

Table 5. Comparison of AI Models for GHI Forecasting

Model	**RMSE (Wh/m^2)**	**MAE (Wh/m^2)**	**R^2**
Random Forest	0.646	0.437	0.834
LSTM	0.794	0.596	0.749
Transformer	0.769	0.518	0.765
CNN	0.813	0.611	0.738
LightGBM	**0.640**	**0.430**	**0.840**

These results confirm that the LightGBM model outperforms the other tested architectures across all evaluation metrics. While deep learning models such as LSTM and Transformer demonstrated potential in capturing temporal and sequential dependencies, they required more extensive training time and were more sensitive to overfitting. The CNN model, though computationally efficient and adept at capturing short-term patterns, struggled to model seasonal and long-range dependencies, resulting in lower accuracy. In contrast, LightGBM combined robust predictive performance with low computational overhead, making it the most suitable candidate for scalable, real-time deployment across Algerian regions.

Wind prediction models are currently under development and will be integrated in future system iterations.

8.4. Interpretation and Insights

- **Hybrid configurations** (particularly 70/30 solar-wind) show the most balanced performance for year-round deployment, mitigating the intermittency of single-source reliance.
- **AI solar forecasting** supports proactive energy management and optimal electrolyzer scheduling by predicting production dips or surges.
- **Geographical adaptation** is essential: Sahara-based systems benefit more from solar-dominant scenarios, while northern/coastal regions may require higher wind integration and potential storage or grid connection.

8.5. AI-Based Wind Forecasting: Performance and Limitations

To complement solar forecasting, we investigated AI-based wind resource prediction for both wind speed (WS50M) and wind direction (WD50M) using three deep learning architectures: LSTM, GRU and Transformer. The objective was to explore model feasibility for future integration into hydrogen yield prediction and optimize hybrid renewable systems.

Table 6. Performance Metrics of Wind Forecasting Models (Unscaled)

Model	Target	MAE	RMSE	R^2
2*Transformer	WS50M	1.343	1.688	0.261
	WD50M	30.604	63.699	0.625
2*LSTM (4-layer)	WS50M	1.358	1.706	0.245
	WD50M	34.072	65.305	0.606
2*GRU (4-layer)	WS50M	1.356	1.704	0.247
	WD50M	33.431	66.411	0.593

Despite reasonable performance in predicting wind direction particularly by the Transformer model (R^2 up to 0.625) none of the models achieved satisfactory accuracy for wind speed estimation. The variability, inconsistency and noisy distribution of wind data across Algerian regions significantly hampered model generalization. These limitations are compounded by sharp discontinuities and seasonal shifts, which introduce disorientation in temporal learning patterns. While the current models are not yet ready for deployment, their evaluation provides critical insights into data preprocessing needs, temporal feature engineering and the suitability of hybrid AI architectures for wind forecasting. Refinement of these components will be essential for the next iteration of the platform. Us as developers and as AI engineers, we're focusing on refining wind forecasting models by addressing data variability and temporal instability, with the goal of integrating real-time wind predictions alongside solar. This will enable a more holistic simulation of hybrid renewable systems. The platform will also be extended to support economic evaluation, smart storage strategies and dynamic operational control.

9. Limitations

This study delivers a functional AI-driven platform for forecasting and simulating green hydrogen production across Algerian regions. However, several limitations should be acknowledged to contextualize the current system's scope and inform its future development.

First, although the platform supports multiple electrolyzer types, the simulation scenarios in this article were performed using only one predefined configuration. This choice was made to ensure comparability across energy mixes and regions. The implications of using alternative electrolyzer technologies especially those with varying efficiency curves and operational constraints - remain unexplored in this version and are left for future iterations.

Second, while the use of NASA POWER data enables globally accessible and scalable implementation, its spatial and temporal resolution may not fully capture the heterogeneity of microclimatic conditions in complex terrains such as Algeria's mountainous and desert zones. This could lead to localized inaccuracies in both forecasting and energy yield simulations, especially where ground-level measurements differ significantly from satellite-derived estimates.

Finally, wind forecasting is still in development. Although preliminary deep learning models for wind speed and direction have been tested, their performance remains insufficient for integration into the real-time production pipeline. As a result, hydrogen yield simulations based on wind inputs rely on historical data rather than AI-predicted values. This restricts the accuracy and applicability of the current hybrid (solar/wind) and wind-only scenario assessments.

Addressing these limitations will require interdisciplinary collaboration across meteorology, process engineering, economics and policy modeling to support more holistic and deployment-ready hydrogen planning tools.

10. Economic feasibility of green hydrogen production in Algeria

The economic viability of Algeria's green hydrogen ambitions rests on its ability to translate abundant renewable resources into cost-competitive hydrogen through efficient infrastructure and effective industrial scaling. This analysis incorporates spatial heterogeneity in resource quality, capital investment requirements, infrastructure constraints and export competitiveness within the evolving global hydrogen economy.

10.1. Levelized Cost of Hydrogen (LCOH): Current Benchmarks and Forecasts

Algeria's southern provinces, where annual solar irradiation exceeds 2,600 kWh/m², offer some of the world's most advantageous conditions for photovoltaic electricity generation. With the levelized cost of electricity (LCOE) projected to fall below \$0.025/kWh by 2030, Algeria is positioned to produce green hydrogen at globally competitive prices. According to IRENA and the Hydrogen Council, under optimal conditions and with mature electrolyzer deployment, Algeria could reach a levelized cost of hydrogen (LCOH) in the range of \$1.5–\$2.0/kg H_2 by 2030, potentially falling below \$1/kg H_2 by 2050 under high- learning-rate scenarios (Hydrogen Council and McKinsey & Company, 2023; International Renewable Energy Agency, 2020). However, Algeria's current installed electrolyzer capacity is under 5 MW, underscoring the need for rapid industrial-sale ramp-up. Empirical regional assessments reinforce this outlook. For example, Tiar et al. (2024) estimate that, assuming continued capital cost reductions and high electrolyzer utilization, production costs in Tamanrasset and Illizi could stabilize around \$1.1-\$1.2/kg H_2 by mid-century. Electricity input is a key variable, accounting for 30-60% of total LCOH according to the Clean Hydrogen Observatory (2023). This underscores Algeria's leverage through low-cost solar PV, though inland sites may face higher OPEX due to water infrastructure limitations and desalination needs.

10.2. Infrastructure and Transport Economics

While domestic production costs are promising, delivered hydrogen prices - especially for export - are highly sensitive to infrastructure and logistics. Algeria's extensive gas export infrastructure, notably the Medgaz and Transmed pipelines, offers retrofitting potential for hydrogen transmission. Retrofitting these pipelines is projected to keep transport costs under \$0.25/kg H_2 by 2030, substantially lower than maritime liquefaction or ammonia reconversion pathways (Drenkard & Mirakyan, 2021). However, regional disparities complicate the cost landscape. GIZ-supported Power-to-X feasibility studies reveal that LCOH can range from \$4.35/kg in optimized southwestern locations (Tindouf, Adrar) to over \$5.50/kg in remote, infrastructure-sparse areas when full logistics costs are internalized (Drenkard & Mirakyan, 2021). These findings highlight the importance of strategic site selection and phased infrastructure deployment to minimize both capital intensity and marginal delivery costs.

10.3. Export Market Competitiveness

Algeria's competitiveness in hydrogen export markets, particularly to Europe, is enhanced by its proximity and existing trade relations. A 2022 comparative study by the Africa Green Hydrogen Alliance and McK- insey ranked Algeria among the top five most cost-effective suppliers of green ammonia to Central Europe, with expected landed costs between \$1.15 and \$1.30/kg H_2 equivalent - figures that compare favorably with projected European domestic production costs (Africa Green Hydrogen Alliance and McKinsey & Company, 2022). Moreover, Algeria's existing ammonia production infrastructure from grey hydrogen may serve as a transitional platform, facilitating early-stage exports of green derivatives and aligning with EU carbon border adjustment mechanism (CBAM) requirements.

10.4. Structural Constraints and Economic Risks

Algeria faces two main structural barriers:

– **High Capital Expenditure:** Electrolyzer CAPEX, especially for PEM and SOEC technologies, remains high despite a 60% cost reduction between 2010 and 2020. Further reductions depend on integration into global value chains and local manufacturing capacity development (Drenkard & Mirakyan, 2021).

– **Water Availability:** Electrolysis requires about 9 liters of deionized water per kg H_2. While Algeria ranks second among Arab countries in desalination capacity, using desalinated seawater adds energy costs and logistical complexity, especially for inland projects. These factors must be explicitly considered in techno-economic models to avoid overestimating production potential (Drenkard & Mirakyan, 2021).

10.5. Policy Alignment and Strategic Vision

Algeria's National Hydrogen Strategy (2023) outlines a three-phase roadmap: technical demonstration (2023-2030), market maturation (2030-2040) and industrial-scale export (2040-2050). The target is to deploy 10 GW of electrolysis capacity by 2040, aiming for 40 TWh annual production, with 75% allocated for export. The strategy favors a 70:30 solar-to-wind mix, reflecting Algeria's solar resource profile and land availability. Institutional reforms, the development of hydrogen standards and certifications and adaptation of technical education are underway to localize industrial capability and reduce dependence on imported technologies (Drenkard & Mirakyan, 2021).

10.6. Infrastructure Strategy

The proposed *SoutH$_2$ Corridor* a 3,300 km hydrogen-dedicated pipeline connecting Algeria to Southern and Central Europe - is central to Algeria's export vision. Its realization would enable Algeria to become a linchpin in Europe's post-gas energy security architecture, leveraging existing gas pipeline corridors and minimizing sunk infrastructure costs (Energy News Pro, 2025).

10.7. International Partnerships and Investment Diplomacy

Algeria has accelerated bilateral and multilateral hydrogen diplomacy, including the Algeria-Germany Joint Declaration (2024) and memoranda with Sonatrach, Sonelgaz and European partners. These initiatives de-risk early-stage investment, facilitate knowledge transfer and promote convergence with European hydrogen certification regimes (Energy News Pro, 2025).

10.8. Strategic Policy Proposition

A robust policy for Algeria should embed green hydrogen in a broader industrial decarbonization strategy, coupling hydrogen production zones with low-carbon steel, ammonia and synthetic fuel industries to create regional value chains, especially in high-unemployment regions (Green Hydrogen Organisation, 2024).

Pillar	Policy Action	Algerian Contextual Implementation
Institutional Governance	Centralized hydrogen regulatory authority with interministerial coordination.	National Green Hydrogen Council reporting to the Prime Minister; coordination among Sonatrach, Sonelgaz, ministries.
Legal & Regulatory Framework	Dedicated hydrogen law, EU-aligned certification, export licensing.	Model on hydrocarbon/renewables law; integrate PPPs; clarify land rights for southern projects.
Infrastructure Deployment	Phased pipeline retrofitting, SoutH$_2$ Corridor, hydrogen blending trials.	Technical feasibility studies with EU partners; begin blending trials on Sonelgaz lines.
Domestic Industrial Integration	Co-location with value-added industries (ammonia, steel, methanol).	Extend industrial zones in Adrar/Tamanrasset; incentivize green fertilizer, synthetic fuel plants.
Investment & Finance	Mobilize international and public-private capital.	SONATRACH as anchor investor; fiscal incentives under new Investment Law; negotiate with international funds.
Human Capital & Training	Specialized programs in hydrogen systems and engineering.	Expand CDER programs; partner with GIZ/IRENA for curriculum and technical training.
Water-Energy Nexus	Solar-powered desalination, circular water use in electrolysis.	Expand desalination capacity; incentivize hybrid solar-desalination units in Adrar/Béchar.
International Market Access	Secure long-term offtake agreements with EU, ensure CBAM compliance.	Formalize bilateral energy cooperation; leverage Sonatrach's trading subsidiaries for pilot contracts.
Regional Energy Integration	North African hydrogen platform, intra-African corridors.	Propose Maghreb Green Hydrogen Cooperation Agreement; co-develop regional pilot projects.
Just Transition & Inclusion	Job creation, gender equity, SME participation in hydrogen value chains.	Job quotas, local content requirements, targeted SME subsidies in southern projects.

Figure 11. Summary Table: Strategic Pillars for Green Hydrogen Development in Algeria

11. Future Work

To advance the platform from a simulation prototype to a fully operational, decision-support system for green hydrogen planning in Algeria, several key directions are envisaged:

- **Expansion Beyond Solar and Wind:** While current simulations focus on solar and wind inputs, future iterations of the platform will integrate other renewable resources, particularly **biomass energy** derived from sustainable forest residues and agro-industrial waste. This aligns with Algeria's forest conservation goals and opens the path for *carbon-neutral hydrogen* from circular bio- resources.

- **Forestry-Energy Nexus:** The role of forests in the hydrogen value chain will be explored not only as biomass sources but also as climate buffers in land-use planning for green infrastructure. Spatial mapping tools will be extended to evaluate forest density, fire risk and reforestation potential in coordination with hydrogen infrastructure deployment.
- **Wind Energy Forecasting Module:** A LightGBM-based model for wind prediction - leveraging wind speed, direction and pressure data - is under development. This will be integrated with the existing system to support dynamic hybrid generation modeling and improve regional scenario planning.
- **Techno-Economic Optimization:** Models incorporating Levelized Cost of Hydrogen (LCOH), electrolyzer CAPEX/OPEX trends and transport/storage costs will be integrated. This will inform investment prioritization and align platform outputs with Algeria's 2040 Hydrogen Roadmap.
- **Grid and Storage Planning:** Reinforcement learning and mixed-integer optimization techniques will be evaluated for smart grid integration, hydrogen buffering and peak shaving strategies. The aim is to minimize curtailment while maintaining infrastructure flexibility.
- **Probabilistic Forecasting and Risk Analysis:** Uncertainty-aware models (e.g., quantile regression forests, Monte Carlo dropout in deep nets) will be tested to provide confidence bands and support resilient energy planning in the face of renewable variability.
- **Blockchain and Certification Frameworks:** To support export readiness, future versions may integrate blockchain-based certification of green hydrogen production - ensuring transparency in CO_2 footprint tracking and compliance with EU and global clean hydrogen standards.
- **User-Centric Platform Deployment:** The platform will evolve toward a web-based interactive tool usable by policymakers, researchers and regional planners. Custom scenario design, regional data downloads and visualization dashboards will be prioritized.

12. Conclusion

This study introduces a modular, AI-driven platform for forecasting and simulating green hydrogen production across Algeria. Leveraging real-time meteorological data from NASA POWER and an integrated electrolyzer simulation environment, the platform enables scenario-based analysis of renewable energy configurations specifically 100% solar, 100% wind and hybrid mixes (70/30 and 50/50). These simulations were deployed across climatically diverse regions to assess spatial performance variability.

Solar forecasting was achieved using a LightGBM regression model, which consistently outperformed deep learning alternatives (LSTM, Transformer, CNN) in predicting Global Horizontal Irradiance (GHI). This enabled precise anticipation of solar-derived hydrogen output under varying conditions. Wind forecasting, in contrast, remains a developmental component. Although recurrent and attention-based architectures (LSTM, GRU, Transformer) were evaluated for wind speed and direction prediction, their performance was hindered by data noise and short-term variability. Consequently, wind scenarios in the current version rely on historical time series.

Simulation outcomes highlighted pronounced regional disparities in green hydrogen potential. Southern provinces such as Adrar and Tamanrasset characterized by high and stable solar irradiance - demonstrated the greatest hydrogen yields. Meanwhile, hybrid energy scenarios provided more consistent performance in semi-arid regions like Djelfa and Batna, where solar and wind resources are seasonally complementary. These findings underscore the strategic value of geographic tailoring in national hydrogen infrastructure planning.

The proposed platform lays the groundwork for a decision-support system that can inform site selection, infrastructure prioritization and energy mix optimization at both local and national scales. By combining AI forecasting with techno-economic simulation, it offers a scalable and adaptable tool aligned with Algeria's 2040 green hydrogen roadmap. Future extensions will include robust wind modeling, dynamic LCOH estimation and integration of smart energy management modules such as battery storage, demand-side response and grid coupling mechanisms. These additions aim to enhance the system's fidelity, usability and relevance for both policy makers and industry stakeholders engaged in the energy transition.

Acknowledgment

This work was conducted with the support of the Faculty of Engineering's research and development team, whose technical input contributed to the implementation of the forecasting and simulation platform. The NASA POWER project is acknowledged for providing open-access meteorological datasets, which constituted a core data source for the system's predictive components.

References

Abed, J., Sargent, E. et al. (2024). Team using AI finds a cheaper way to make green hydrogen. https://phys.org./

Africa Green Hydrogen Alliance and McKinsey & Company (2022). Africa's green hydrogen potential. https://gh2.org/initiatives/africa-green-hydrogen-alliance Accessed on 07.06.2025.

Alhussan, A.A., El-Kenawy, E.S.M., Saeed, M.A., Ibrahim, A., Abdelhamid, A.A., Eid, M.M., ... & Elbaksawi, O. (2023). Green hydrogen production ensemble forecasting based on hybrid dynamic optimization algorithm. *Frontiers in Energy Research*, 11, 1221006.

Benchenina, Y., Zemmit, A., Bouzaki, M.M., Loukriz, A., Elsayed, S.K., Alzaed, A., ... & Ghoneim, S.S. (2025). Advancing green hydrogen production in Algeria with opportunities and challenges for future directions. *Scientific Reports*, *15*(1), 5559.

Borovykh, A., Bohte, S. & Oosterlee, C.W. (2017). Conditional time series forecasting with convolutional neural networks. *arXiv*, arXiv:1703.04691. Accessed on 06.06.2025.

Bosma, S.B., Nazari, N. (2021). Estimating solar and wind power production using computer vision deep learning techniques on weather maps. *arXiv*, arXiv:2103.08727. https://arxiv.org/abs/2103.08727.

Bouchair, A. et al. (2023). Wind energy potential in Algeria: A review. https://www.researchgate.net/publication/330203275 Accessed on 07.06.2025.

Drenkard, S., Mirakyan, A. (2021). Exploratory study on the potential of power-to-x (green hydrogen) for Algeria. Deutsche Gesellschaft fu¨ r Internationale Zusammenarbeit (GIZ). Accessed on 07.06.2025.

Energy Box (2025). Algeria launches renewable energy, hydrogen plan. https://www.energy-box.com/post/algeria-launches-renewable-energy-hydrogen-plan Accessed on 06.06.2025.
Energy News Pro (2025). Algeria, pivot of the hydrogen corridor to Europe. https://energynews.pro/en/algeria-pivot-of-the-hydrogen-corridor-to-europe/ Accessed on 07.06.2025.
Ensure IAS (2024). Electrolysers: A key component in green hydrogen production. Accessed on 05.06.2025.
European Hydrogen Observatory (2023). Levelised cost of hydrogen calculator. https://observatory.clean-hydrogen.europa.eu/tools-reports/levelised-cost-hydrogen-calculator Accessed on 07.06.2025.
Gougui, A., Djafour, A., Danoune, M.B. & Khelfaoui, N. (2019). Field experience study and evaluation for hydrogen production through a photovoltaic system in Ouargla region, Algeria. *International Journal of Hydrogen Energy*, *45*(4), 2593-2606.
Green Hydrogen Organisation (2024). Algeria. https://gh2.org/countries/algeria Accessed on 07.06.2025.
Hydrogen Central (2025). Algeria reaffirms commitment to green hydrogen development through southh2 corridor project. https://hydrogen-central.com/algeria-reaffirms-commitment-to-green-hydrogen-development-through-southh2-corridor-project/ Accessed on 06.06.2025.
Hydrogen Council and McKinsey & Company (2023). Hydrogen insights 2023. https://hydrogencouncil.com/en/hydrogen-insights-2023/ Accessed on 07.06.2025.
Institute for Sustainable Alternatives (2023). Algeria's national hydrogen strategy 2023. https://www.isa-ghic.org/national-hydrogen-strategy-2023 Accessed on 07.06.2025.
International Energy Agency (2022). Electrolyzers: Technology brief. https://www.iea.org/reports/electrolysers-technology-brief Accessed on 06.06.2025.
International Energy Agency (2023). Tracking hydrogen 2023: Analysis and key findings. https://www.iea.org/reports/tracking-hydrogen-2023 Accessed on 06.06.2025.
International Energy Agency (2021). Solar pv global supply chains. https://www.iea.org/reports/solar-pv-global-supply-chains Accessed on 06.06.2025.
International Renewable Energy Agency (2020). Green hydrogen: A guide to policy making. https://www.irena.org/publications/2020/Nov/Green-hydrogen-A-guide-to-policy-making Accessed on 07.06.2025.
International Solar Alliance (2023). Algeria - green hydrogen innovation centre. https://isaghic.org/countries/algeria Accessed on 06.06.2025.
K E, V., Padhan, P. (2024). Data-driven catalyst design: A machine learning approach to predicting electrocatalytic performance in hydrogen evolution and oxygen evolution reactions. https://arxiv.org/abs/2412.12846 Accessed on 06.06.2025.
Kumar, A., Singh, R. (2023). Computational methods, artificial intelligence, modeling and simulation for green hydrogen technologies. *Energies*, *6*(2), 21. Accessed on 06.06.2025.
Lai, N.S., Tew, Y.S., Zhong, X., Yin, J., Li, J., Yan, B. & Wang, X. (2024). Artificial intelligence (AI) workflow for catalyst design and optimization. *arxiv*, arXiv:2402.04557.
Lara-Ben´ıtez, P., Carranza-Garc´ıa, M. & Riquelme, J.C. (2021). An experimental review on deep learning architectures for time series forecasting. *arXiv*, arXiv:2103.12057. Accessed on 06.06.2025.
Lesniak, A., Johnsen, A.G., Rhodes, N. & Roald, L. (2024). Advanced scheduling of electrolyzer modules for grid flexibility. https://arxiv.org/abs/2412.19345 Accessed on 06.06.2025.
Li, Y., Wang, R., Li, Y., Zhang, M. & Long, C. (2022). Wind power forecasting considering data privacy protection: A federated deep reinforcement learning approach. https://arxiv.org/abs/2211.02674
Liao, W., Bak-Jensen, B., Pillai, J.R., Yang, Z. & Liu, K. (2021). Short-term power prediction for renewable energy using hybrid graph convolutional network and long short-term memory approach. *arXiv*, arXiv:2111.07958. https://arxiv.org/abs/2111.07958
Nikulins, A., Sudars, K., Edelmers, E., Namatevs, I., Ozols, K., Komasilovs, V., ... & Reinhardt,

A. (2024). Deep learning for wind and solar energy forecasting in hydrogen production. *Energies*, *17*(5), 1053.

Tiar, B., Fadlallah, S.O., Benhadji Serradj, D.E., Graham, P. & Aagela, H. (2024). Navigating Algeria towards a sustainable green hydrogen future to empower North Africa and Europe's clean hydrogen transition. *International Journal of Hydrogen Energy*, 61, 783-802.

Veettil Muhammed, S.N. (2023). How AI can accelerate the transition to green hydrogen. https://blog.se.com/sustainability/2023/07/18/how-ai-can-accelerate-the-transition-to-green-hydrogen Accessed on 06.06.2025.

Wayo, D.D.K., Goliatt, L. & Ganji, D. (2024). AI and quantum computing in binary photocatalytic hydrogen production. *arXiv*, arXiv:2501.00575.

Wikipedia contributors (2024). Solarpower in Algeria. https://en.wikipedia.org/wiki/SolarpowerinAlgeria Accessed on 07.06.2025.

Yin, G., Zhu, H., Chen, S., Li, T., Wu, C., Jia, S., … & Li, Y. (2025). Machine learning-assisted high-throughput screening for electrocatalytic hydrogen evolution reaction. *Molecules*, *30*(4), 759.