Literature review on forecasting green hydrogen production using machine learning and deep learning

Mohamed Yassine Rhafes[1], Omar Moussaoui[1], Maria Simona Raboaca[2]
[1]MATSI Laboratory, High School of Technology (ESTO), Mohammed First University, Oujda, Morocco
[2]Department of ICSI Energy, National Research and Development Institute for Cryogenics and Iso-topic Technologies, Râmnicu Vâlcea, Romania

Article Info

Article history:

Received Apr 14, 2024
Revised Nov 14, 2024
Accepted Nov 24, 2024

Keywords:

Deep learning
Forecasting
Global energy transition
Green hydrogen production
Machine learning

ABSTRACT

Green hydrogen is a sustainable and clean energy source, for this purpose, it conducts the global energy transition. The integration of artificial intelligence (AI), especially machine learning (ML) and deep learning (DL) with the process of green hydrogen production is essential in enhancing its production. This literature review studies in detail the intersection between AI and green hydrogen. Firstly, it concentrates on ML and DL algorithms used in forecasting green hydrogen production. Secondly, it presents an analysis of the studies released from 2021 to March 2024. Finally, the focus is on the results realized by the ML and DL algorithms proposed by the studies reviewed. This study provides a summary that explains the trends and methods used, as well as highlights the gaps and the opportunities in the field of AI and green hydrogen production. This liternature review presents a solid foundation for future research initiatives in this field.

Corresponding Author:

Mohamed Yassine Rhafes
MATSI Laboratory, ESTO, Mohammed First University
Oujda, Morocco
Email: mohamedyassine.rhafes@ump.ac.ma

1. INTRODUCTION

Renewable energy sources such as solar energy, wind energy, and others are becoming an important option in the energy sector [1], the process of producing green hydrogen starts with these sources that generate green electricity. Green hydrogen become an alternative to traditional fossil fuels due to the cleaner process of producing energy. The advantage of green hydrogen is that when it's used, it only produces water as a byproduct, unlike fossil fuels that produce dangerous gas emission [2], this quality makes it as an important element to build greener and sustainable future. Another important aspect is that green hydrogen can be stored for long periods with little energy loss [3] makes it a long-term energy solution.

Producing green hydrogen with electricity, especially using renewable energy sources, is a greener choice compared to old ways of making hydrogen. This fits with worldwide goals to use less fossil fuels and protect the environment. A great use of green hydrogen is to power electric cars [4], which are cleaner and more efficient than traditional vehicles. Other uses include power generation, heating, and various other applications [5]. To produce green hydrogen, the following three steps are essential as described in Figure 1. Renewable energy: the process begins with the generation of electricity from renewable energy sources [6], [7] for example, solar panels [8], wind turbines [9], or hydroelectric plants [10]. The electricity needs to come from renewable energy sources to confirm that hydrogen production is sustainable and does not emit greenhouse gases. Electrolysis process [11]–[14]: electrolysis involves splitting water (H_2O) into its basic components, hydrogen (H_2) and oxygen (O_2). This is achieved by applying an electrical current to water that has an electrolyte added to it, which helps in the conduction of electricity. The hydrogen gas collects at the cathode (the negative electrode), and oxygen gas collects at the anode (the positive electrode). Green hydrogen output: the result is green hydrogen [15], [16] that can be used in various applications [5].

However, the passage to universal use of green hydrogen comes with challenges. Key issues include the unpredictability of weather, varying conditions across different areas, and the complexity of modeling energy systems [17]. To reduce these challenges and make green hydrogen more practical, we need to prepare forecasts in advance. In recent years, artificial intelligence (AI), particularly machine learning (ML), and deep learning (DL), has emerged as a tool in addressing these challenges within the renewable energy sector. ML and DL models are adept at processing and learning from vast datasets, including time series, meteorological, and geographical data. This capability is crucial for developing predictive models for green hydrogen production. The novelty of this work lies in its comprehensive presentation of a state-of-art for future research on forecasting green hydrogen production. By synthesizing current methodologies and trends in the field of ML and DL. We didn't find any literature review papers specifically addressing forecasting green hydrogen production, highlighting the uniqueness and importance of our study in setting the groundwork for future investigations in this area.

The structure of this paper is as follows: section 2 explores in detail statistical methods, ML, and DL algorithms, and performance metrics that are commonly used in the field of forecasting green hydrogen production. Section 3 presents the methodology of our survey, detailing the strategies and criteria employed for selecting and analyzing relevant and comprehensive articles related to the subject. Section 4 combines results and discussions for a clearer understanding. The paper concludes by summarizing the main findings and implications derived from this survey.

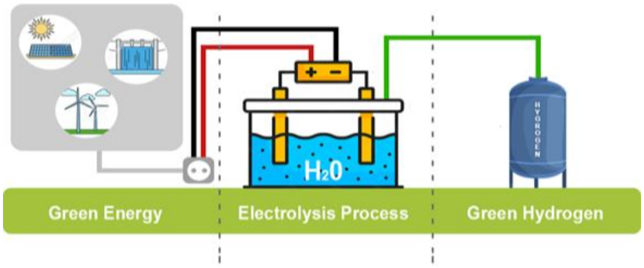


Figure 1. Green hydrogen production pipeline

2. BACKGROUND

2.1. Machine learning, deep learning, and statistical methods in hydrogen production forecasting

This section examines the statistical methods, ML and DL models used in the papers reviewed in section 3. Table 1 displays the statistical methods. Table 2 outlines the ML and DL models.

Table 1. Statistical methods used in the selected studies

Ref	Statistical method	Short description
[18]	Arithmetic optimization algorithm (AOA)	An innovative optimization technique inspired by fundamental arithmetic operations: addition, subtraction, multiplication, and division. It optimizes model hyperparameters, enhances feature selection, and boosts the overall algorithm performance.
[19], [20]	Complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN)	A sophisticated signal processing technique that builds upon the empirical mode decomposition (EMD) method. It decomposes complex signals into simpler, constituent components known as intrinsic mode functions (IMFs). This method is especially valuable for feature extraction and noise reduction in time series data. By breaking down signals into their simpler components, CEEMDAN enables algorithms to more effectively identify underlying patterns and trends, thereby enhancing predictive accuracy.
[21], [22]	Al-Biruni earth radius (BER)	A historical method was created by the Persian scholar Al-Biruni to determine the Earth's radius. This innovative approach involves observing the horizon from a high vantage point and applying geometric principles and trigonometry to estimate the Earth's curvature and, subsequently, its radius.
[23], [24]	Particle swarm optimization (PSO)	An optimization technique inspired by the social behavior of birds and fish. It optimizes model hyperparameters, enhances feature selection, and improves overall algorithm performance. This is achieved by simulating the collective movement of a group of points, referred to as particles, through the solution space. The direction of each particle is adjusted based on both individual discoveries and the group's collective findings of optimal solutions.

Table 2. ML and DL models used in the selected studies

Ref	ML/DL model	Short description
[25]–[27]	Extreme learning machine (ELM)	A feedforward neural network characterized by the presence of one hidden layer with randomly allocated weights offers substantial benefits, including accelerated learning speeds and simplified implementation.
[28], [29]	Convolutional neural network (CNN)	A neural network optimized, primarily designed to process and interpret data with a grid-like topology. CNNs can handle time-series data for forecasting and anomaly detection in various fields like weather prediction.
[30], [31]	Recurrent neural network (RNN)	A type of neural network designed for sequential data, such as time-series data. It's characterized by its ability to maintain an internal memory of previous inputs in a sequence, allowing it to capture temporal dependencies and context in data.
[30], [32]	Long short-term memory (LSTM)	An advanced version of RNNs, specialized in remembering information for extended periods. It's highly effective in complex sequence prediction tasks like time series analysis, where long-term context is crucial.
[33]	Gated recurrent unit (GRU)	A type of RNN that effectively captures dependencies in sequences, with a simpler structure than LSTM, improving efficiency.
[34]	Categorical boosting (CatBoost)	A ML algorithm optimized for supervised learning tasks. It's designed to provide high accuracy while requiring minimal data preprocessing and parameter tuning.
[35], [36]	Support vector machine (SVM)	A ML algorithm is used for classification and regression. It works by finding the hyperplane, that separates data points by the largest margin possible.
[37]	Prophet	A forecasting tool designed for handling time series data. It's particularly effective for data with strong seasonal effects and several seasons of historical data. It works well with daily observations and can handle missing data and trend changes.
[38], [39]	Linear regression (LR)	A ML algorithm is used in regression problems. It models the relationship between a dependent variable and independent variables.
[40]	Seasonal auto regressive integrated moving average exogenous (SARIMAX)	An advanced statistical model used for forecasting time series data. SARIMAX extends the ARIMA model by incorporating seasonal trends and external (exogenous) variables, making it highly effective for complex time series with seasonal patterns.
[41], [42]	Stochastic gradient descent regression (SGDR)	A ML algorithm is used for regression tasks. It employs the SGD optimization method, making it efficient for large datasets.

2.2. Measurements of forecasting performance

Performance measurements consist of a variety of statistical tools and methodologies employed to evaluate and quantify the effectiveness of a model [43]. Table 3 illustrates the performance metrics used in selected research studies, detailing the metrics, their formulas, and the components involved in each formula. This helps in directly comparing the quantitative aspects of different regression models.

Table 3. Performance metrics used in the selected studies

Metric	Formula	Components
Mean absolute error (MAE)	$\frac{1}{N}\sum_{i=1}^{N}\lvert y_i - \hat{y}_i\rvert$	N is the number of observations. y_i is the actual value. $\hat{y}_i$ is the predicted value. $\bar{y}_i$ the mean of the actual values.
Mean squared error (MSE)	$\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$	
Root mean squared error (RMSE)	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$	
Mean absolute percentage error (MAPE)	$\frac{100}{N}\sum_{i=1}^{N}\frac{\lvert y_i - \hat{y}_i\rvert}{y_i}$	
R^2 score (coefficient of determination)	$1 - \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}}$	
Standard deviation	$\sqrt{\sum_{i=1}^{N}\frac{1}{N-1}(y_i - \hat{y}_i)^2}$	

2.3. Forecasting horizons

Forecasting horizons refer to the time periods for which predictions are made. The accuracy of ML and DL algorithms in predicting green hydrogen production from renewable energy sources is dependent on the time horizon of the forecasts [44]. Table 4 presents the categories of forecasting horizons.

Table 4. Categories of forecasting horizons [44]

Type	Description
Short-term	few minutes or hours up to 72 hours ahead
Medium-term	from around 72 hours to a few weeks ahead
Long-term	from several weeks to several months or even years ahead

3. METHODOLOGY

In our study, we focused on English language papers and selected the Scopus database to find articles, as it's known for offering high-quality data. Scopus includes a diverse array of publications, such as journal and conference papers, patents, and various websites in significant fields, as cited in the source [45]. Figure 2 presents the papers search process used in our literature review. We used the search terms [("forecasting" OR "prediction") AND ("green hydrogen production" OR "hydrogen production")], and found approximately 1019 documents, indicating significant research interest in hydrogen production. To narrow our focus, we refined our search using [("forecasting" OR "prediction") AND ("green hydrogen production" OR ("hydrogen production" AND ("renewable energies" OR "green energy")) AND ("machine learning" OR "deep learning"))], relevant to our review topic, and found about 25 documents. These papers were published between 2018 and March 2024, as presented in Figure 3. For our analysis, we carefully chose articles that specifically focused on green hydrogen production, particularly emphasizing their connection with ML and DL algorithms. Table 5 presents the papers selected for this study.

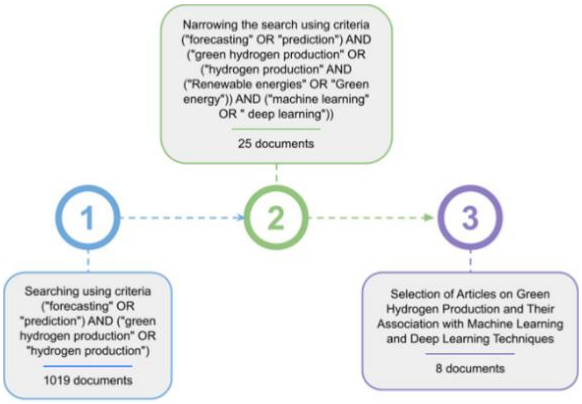


Figure 2. Papers search process

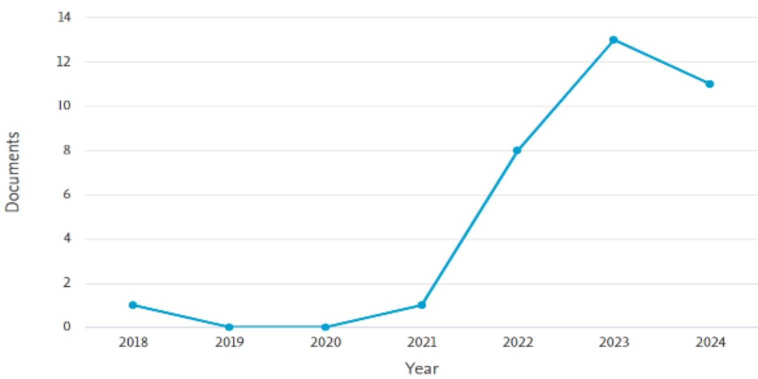


Figure 3. Scopus indexed papers per year using the terms [("forecasting" OR "prediction") AND ("green hydrogen production" OR ("hydrogen production" AND ("renewable energies" OR "green energy")) AND ("machine learning" OR " deep learning"))]

Table 5. Papers selected for review

Year	Article	Short description
2024	[46]	The study advances the field of hydrogen production powered by wind and solar energy through the creation of advanced DL models for weather prediction. Using FCN and CNN models, it forecasts weather patterns essential for generating hydrogen from renewable energy sources. Leveraging data from 25 weather stations across Latvia, the models achieve significant accuracy in predicting energy outputs.
2024	[47]	The study presents a green hydrogen production method using solar energy, optimized by ML. It evaluates four ML models for predicting hydrogen output from a solar-powered system, with the CatBoost model optimized by AOA showing the best accuracy.
2024	[48]	The study discusses improving green hydrogen production from solar energy by addressing the variability in solar power output, which impacts the consistent electricity supply needed for electrolysis. The study introduces a forecasting algorithm CEEMDAN-bidirectional long short-term memory (BiDLSTM), to predict solar-based hydrogen production potential accurately.
2024	[49]	The study focused on enhancing green hydrogen production forecasting using CEEMDAN-GRU. with the energy for electrolysis sourced from wind energy. It introduces a novel method for determining green hydrogen production by analyzing the wind patterns across 9 selected provinces, identifying the most suitable wind turbine power based on wind speed decomposition.
2023	[50]	The study introduces a forecasting method for solar hydrogen generation using an optimized RNN model, combining BER and PSO. This method aims to enhance the accuracy and efficiency of predicting solar energy production, showing promising results in comparison to existing forecasting techniques. The effectiveness of the BER-PSO-RNN algorithm is validated by statistical tests, demonstrating its potential in optimizing the operation of sustainable energy systems.
2023	[51]	This study evaluates the potential of green hydrogen production via photovoltaic-powered water electrolysis in China. By forecasting green hydrogen production using a photovoltaic-electrolysis system and ML methods, it finds the non-time series algorithm SVM outperforms FbProphet in accuracy. Results show high R^2 values and varying RMSE across four regions, with significant daily production potential in high radiation areas.
2022	[52]	This study investigates the potential of employing wind energy for green hydrogen production in a suburban environment. It explores the use of AI techniques, particularly LSTM, support vector regression (SVR), and LR algorithms, to predict daily green hydrogen output based on wind data collected at a specific site. The findings suggest that LSTM models perform best in predicting green hydrogen production, demonstrating the feasibility of this renewable energy approach.
2021	[53]	This study explores forecasting the solar hydrogen production potential in Islamabad, Pakistan, using ML algorithms: Prophet, SARIMAX, and SGDR. Focusing on a photovoltaic-electrolytic system. Among the three algorithms tested, Prophet was the most accurate, especially for the months transitioning into winter.

4. RESULTS AND DISCUSSION

In this section, we elaborate on the papers selected from the previous section. Firstly, we outline the criteria used for comparing these results as presented in Table 6. i) Models: models used in the mentioned paper to forecast green hydrogen production; ii) Dataset: dataset used to train, test, and evaluate the proposed model; iii) Features: inputs used to train the proposed model; iv) Targets: outputs of the proposed model; v) Metrics: methods used to evaluate and optimize the performance of the model; and vi) Best model: the model that delivered the best performance.

V1: wind speed, V2: temperature, V3: wind direction, V4: precipitation, V5: atmospheric pressure, V6: relative humidity, V7: snow depth, V8: visibilty meteorological, V9: solar irradiance, V10: timestamp, V11: global horizontal irradiance, V12: sunshine hours, V13: fixed month, V14: wind gust, V15: diffuse horizontal irradiance, V16: direct normal irradiance, V17: difference between measure and calculated diffuse horizontal irradiance. The results presented in Table 6 indicate that the studies employ two categories of algorithms within ML and DL, hybrid models at 41.2% and standard models at 58.8%. Figure 4 illustrates the distribution for each type of algorithm used in the studies.

An important factor is the variability of algorithms used. Neural network (NN) models are the most used, at 52.9% of the total studies reviewed. They are followed by SVM, Prophet, and linear regression models (LR, SGDR), each at 11.8%. CatBoost and SARIMAX are each at 5.9%. This explains that neural network models can model complex non-linear relationships within large datasets compared to other models. Figure 5 illustrates the usage of the models in the studies.

According to Javaid *et al.* [52], LSTM, SVM, and LR are used in predicting wind direction, with LSTM outperforming SVM and LR. This underscores the potential of NN in capturing patterns over time, demonstrating that NNs can outperform traditional ML algorithms in certain scenarios. In opposition the study [47] combines statistical methods and standard models, specifically AOA-ELM, AOA-CNN, AOA-GRU, and AOA-CatBoost, and demonstrates that AOA-CatBoost outperforms NN algorithms. We conclude that it is important to select the appropriate model based on the specific characteristics and requirements of the problem.

Table 6. Overview of predictive models in the selected papers

Year	Article	Models	Dataset	Features	Targets	Metrics	Best model	Horizon
2024	[46]	FCN, CNN	Latvian open data portal in 25 weather stations	V1, V2, V3, V4, V5, V6, V7, V8, V10	Wind speed	MSE	CNN	Short term
2024	[47]	AOA-ELM, AOA-CNN, AOA-GRU, AOA-CatBoost	Numerical simulations	V1, V5, V9, V10	Electrical power, thermal power, electrical efficiency, thermal efficiency, hydrogen production	MSE, MAE, R2-score	AOA-CatBoost	Short term
2024	[48]	CEEMDAN-BiDLSTM	National portal of the National Institute of Wind Energy in India	V9, V10	Solar irrradiance	RMSE, MAE, R2-score	CEEMDAN-BiDLSTM	Short term
2024	[49]	CEEMDAN-GRU	MERRA (Modern Era Retrospective-analysis for Research and Applications - NASA)	V1, V3, V5, V10	Wind speed	RMSE, MAPE, R2-score	CEEMDAN-GRU	Short-term
2023	[50]	BER-PSO-RNN	Meteorological data from the HI-SEAS weather station in Hawaii	V2, V5, V6, V9, V10	Solar energy	MAE, MAPE, RMSE, R2-score	BER-PSO-RNN	Medium-term and Long-term
2023	[51]	SVM, Prophet	National Meteorological Information Center of China Meteorological Administration, rp5.ru weather database, Xihe Energy Big Data Platform, and others	V2, V3, V4, V6, V10 (for Prohpet), V12, V13 (for SVM)	Hydrogen production	MSE, RMSE, R2-score	SVM	Short-tem
2022	[52]	LSTM, SVM, LR	Collected in Pakistan at latitude 33.64° N, longitude 72.98° E, and elevation 500 meters above mean sea level	V1, V3, V10, V14	Wind speed	MAE, Standard deviation	LSTM	Short-term
2021	[53]	Prophet, SARIMAX, SGDR	Custom Weather Variables Dataset from Islamabad at latitude 33.64 N, longitude 72.98 E, 500 meters above mean sea level	V1, V2, V3, V5, V6, V10, V11, V15, V16, V17	Hydrogen production	MAE, MSE, RMSE, MAPE, R2-score	Prophet	Short-term

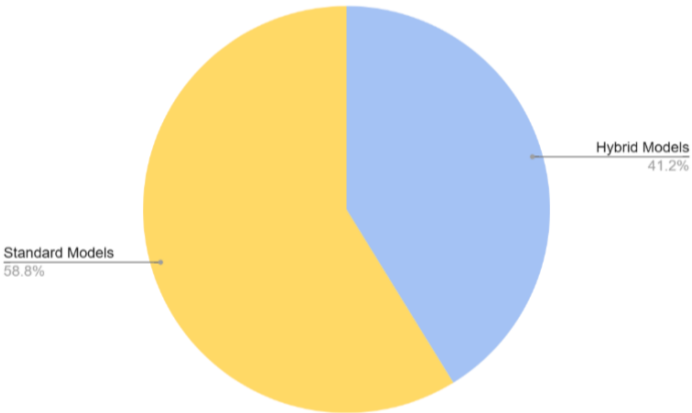


Figure 4. Distribution of model types

According to Cheng *et al.* [51], the SVM outperforms Prophet in predicting hydrogen production. Conversely, the success of the Prophet model over SARIMAX and SGDR in forecasting hydrogen production is demonstrated in study [53]. This highlights the potential of automated forecasting models in effectively handling the seasonal and trend components of time series data. The studies [46], [48]–[50] used NN to predict renewable energy targets in order to examine hydrogen production. This underscores the effectiveness of NN in the renewable energy sector.

In the realm of model evaluation, the R^2-score is the most used, at 25%, followed by RMSE and MAE, each at 20.8%, MSE at 16.7%, MAPE at 12.5%, and standard deviation at 4.2%. The variability indicates that no single metric is universal. Figure 6 illustrates the distribution of performance metrics used in the studies.

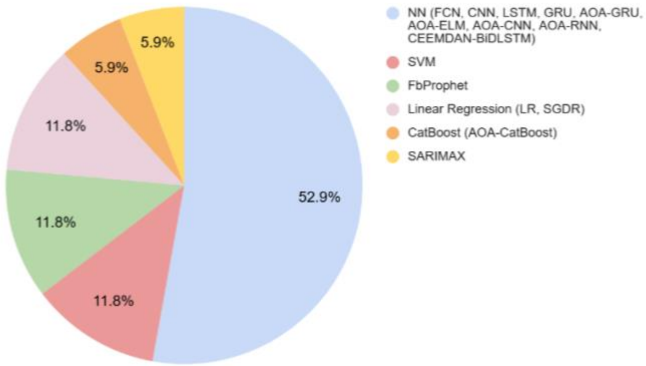


Figure 5. Comparative usage of forecasting models

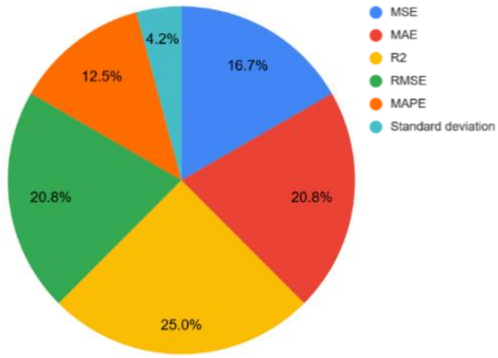


Figure 6. Distribution of performance metrics

In the context of targeting, the studies [46], [48]–[50], [52] employed renewable energy targets to examine hydrogen production. In opposition, the studies [47], [51], [53] directly focused on hydrogen production targets for their investigations. This distinction demonstrates that there is a gap in research focusing only on forecasting green hydrogen production. In the same context, the datasets used in these studies differ in several aspects, the location of the study, the methodology used to gather the data, and the features selected for training, validating, and testing the models.

We found that the feature V10 (timestamp) is present in all datasets [46]–[53]. This is because the objective of the studies is to predict renewable energy targets or green hydrogen production in the future, taking time into account. Additionally, the features V1 (wind speed), V3 (wind direction), and V5 (atmospheric pressure) are used in five studies. Features V2 (temperature) and V6 (relative humidity) are used in four studies. V9 (solar irradiance) is used in three studies, V4 (precipitation) is used in two studies, and the rest of the features are used only in one study. Figure 7 present the distribution of features used in the studies. We conclude that the important features are timestamp, wind speed, wind direction, atmospheric pressure, temperature, relative humidity, solar irradiance, and precipitation. This leads us to research feature selection for examining the forecasting of hydrogen production.

In the realm of forecasting horizons, studies [46]–[53] focus on short-term forecasting, while only study [50] focuses on both medium-term and long-term forecasting. These results highlight the gap in medium-term and long-term forecasting research. Figure 8 illustrates the distribution of forecasting horizons across the studies.

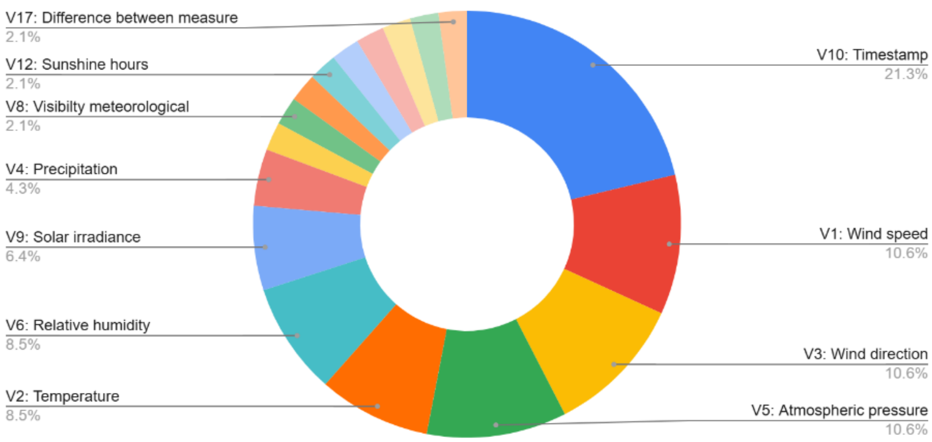


Figure 7. Distribution of features

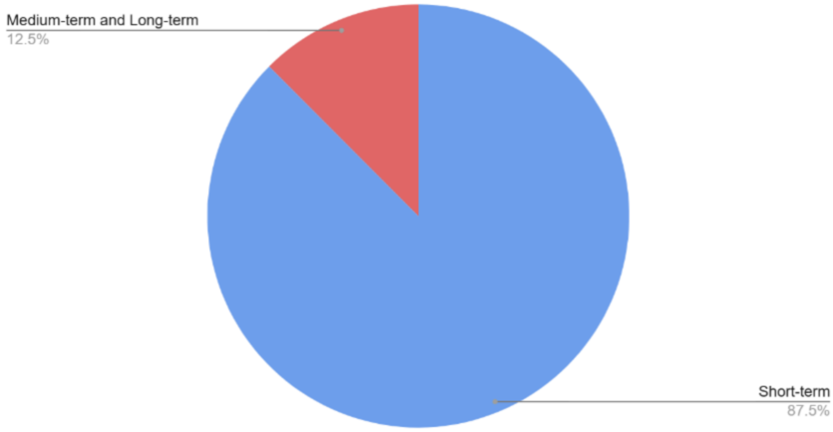


Figure 8. Distribution of forecasting horizons across the studies

5. CONCLUSION

This literature review demonstrated the role of ML and DL in the realm of green hydrogen production. By analyzing studies from 2021 to March 2024, we have captured recent advancements and applications of ML and DL for forecasting green hydrogen production, as well as identified the gaps and opportunities. The results highlighted the algorithms of ML and DL and performance metrics that are most commonly used, the key features and forecasting horizons in the field of forecasting green hydrogen production. For future research initiatives, it is important to carefully select algorithms based on the specific characteristics of the study, this is crucial when focusing on neural networks, which are effective at discovering hidden patterns in time series data. Additionally, the focus on automated forecasting models that demonstrated potential in efficiently managing the seasonal and trend components inherent in such data. Futhermore the importance of feature selection techniques to improve the forecasting accuracy of hydrogen production cannot be overstated. Moreover, a significant research gap exists in medium-term and long-term forecasting, which requires efforts to improve the methods for these time scales.

ACKNOWLEDGEMENT

This work was supported by a grant from the European Regional Development Fund, within the Competitiveness Operational Program, in the frame of project no. 308/2020, HyRo 2.0 project, SMIS 127318 and by the National Plan for Research and Development, PN 23 15 01 01 project.

REFERENCES

[1] M. Ball and M. Weeda, “The hydrogen economy – vision or reality?” *International Journal of Hydrogen Energy*, vol. 40, no. 25, pp. 7903–7919, 2015, doi: 10.1016/j.ijhydene.2015.04.032.
[2] A. Mostafaeipour *et al.*, “Evaluating the wind energy potential for hydrogen production: a case study,” *International Journal of Hydrogen Energy*, vol. 41, no. 15, pp. 6200–6210, 2016, doi: 10.1016/j.ijhydene.2016.03.038.
[3] B. Mohamed, B. Ali, B. Ahmed, B. Ahmed, L. Salah, and D. Rachid, “Study of hydrogen production by solar energy as tool of storing and utilization renewable energy for the desert areas,” *International Journal of Hydrogen Energy*, vol. 41, no. 45, pp. 20788–20806, 2016, doi: 10.1016/j.ijhydene.2016.07.034.
[4] T. Wilberforce *et al.*, “Developments of electric cars and fuel cell hydrogen electric cars,” *International Journal of Hydrogen Energy*, vol. 42, no. 40, pp. 25695–25734, 2017, doi: 10.1016/j.ijhydene.2017.07.054.
[5] H. Ishaq, I. Dincer, and C. Crawford, “A review on hydrogen production and utilization: challenges and opportunities,” *International Journal of Hydrogen Energy*, vol. 47, no. 62, pp. 26238–26264, 2022, doi: 10.1016/j.ijhydene.2021.11.149.
[6] M. Child, C. Kemfert, D. Bogdanov, and C. Breyer, “Flexible electricity generation, grid exchange and storage for the transition to a 100% renewable energy system in europe,” *Renewable Energy*, vol. 139, pp. 80–101, 2019, doi: 10.1016/j.renene.2019.02.077.
[7] L. Gan, G. S. Eskeland, and H. H. Kolshus, “Green electricity market development: lessons from Europe and the US,” *Energy Policy*, vol. 35, no. 1, pp. 144–155, 2007, doi: 10.1016/j.enpol.2005.10.008.
[8] D. S. Vijayan *et al.*, “Advancements in solar panel technology in civil engineering for revolutionizing renewable energy solutions—a review,” *Energies*, vol. 16, no. 18, 2023, doi: 10.3390/en16186579.
[9] M. M. Savino, R. Manzini, V. D. Selva, and R. Accorsi, “A new model for environmental and economic evaluation of renewable energy systems: the case of wind turbines,” *Applied Energy*, vol. 189, pp. 739–752, 2017, doi: 10.1016/j.apenergy.2016.11.124.
[10] C. P. Nazir, “Offshore hydroelectric plant: a techno-economic analysis of a renewable energy source,” *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 174–184, 2014, doi: 10.1016/j.rser.2014.03.006.
[11] S. Sunaryo, “Hydrogen production as alternative energy through the electrolysis process of sea water originating from mangrove plant areas,” *Journal of Physics: Conference Series*, vol. 2377, no. 1, 2022, doi: 10.1088/1742-6596/2377/1/012056.

[12] M. Asare-Addo, "Green hydrogen potential assessment in ghana: application of pem electrolysis process and geospatial-multi-criteria approach," *International Journal of Sustainable Energy*, vol. 42, no. 1, pp. 1202–1225, 2023, doi: 10.1080/14786451.2023.2256892.

[13] D. S. Falcão and A. M. F. R. Pinto, "A review on pem electrolyzer modelling: guidelines for beginners," *Journal of Cleaner Production*, vol. 261, 2020, doi: 10.1016/j.jclepro.2020.121184.

[14] O. Majeed Butt *et al.*, "A comparative study based on performance and techno-economic analysis of different strategies for PV-electrolyzer (green) hydrogen fueling incinerator system," *Waste Management*, vol. 156, pp. 1–11, 2023, doi: 10.1016/j.wasman.2022.11.016.

[15] N. AbouSeada and T. M. Hatem, "Climate action: prospects of green hydrogen in Africa," *Energy Reports*, vol. 8, pp. 3873–3890, 2022, doi: 10.1016/j.egyr.2022.02.225.

[16] R. Raman, V. K. Nair, V. Prakash, A. Patwardhan, and P. Nedungadi, "Green-hydrogen research: what have we achieved, and where are we going? bibliometrics analysis," *Energy Reports*, vol. 8, pp. 9242–9260, 2022, doi: 10.1016/j.egyr.2022.07.058.

[17] G. A. Tiruye, A. T. Besha, Y. S. Mekonnen, N. E. Benti, G. A. Gebreslase, and R. A. Tufa, "Opportunities and challenges of renewable energy production in Ethiopia," *Sustainability*, vol. 13, no. 18, 2021, doi: 10.3390/su131810381.

[18] L. Abualigah, A. Diabat, S. Mirjalili, M. A. Elaziz, and A. H. Gandomi, "The arithmetic optimization algorithm," *Computer Methods in Applied Mechanics and Engineering*, vol. 376, 2021, doi: 10.1016/j.cma.2020.113609.

[19] T. Liang, G. Xie, S. Fan, and Z. Meng, "A combined model based on ceemdan, permutation entropy, gated recurrent unit network, and an improved bat algorithm for wind speed forecasting," *IEEE Access*, vol. 8, pp. 165612–165630, 2020, doi: 10.1109/ACCESS.2020.3022872.

[20] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2011, pp. 4144–4147, doi: 10.1109/ICASSP.2011.5947265.

[21] E.-S. M. El-kenawy *et al.*, "Al-biruni earth radius (BER) metaheuristic search optimization algorithm," *Computer Systems Science and Engineering*, vol. 45, no. 2, pp. 1917–1934, 2023, doi: 10.32604/csse.2023.032497.

[22] A. A. Alghamdi, A. Ibrahim, E.-S. M. El-Kenawy, and A. A. Abdelhamid, "Renewable energy forecasting based on stacking ensemble model and al-biruni earth radius optimization algorithm," *Energies*, vol. 16, no. 3, 2023, doi: 10.3390/en16031370.

[23] F. Marini and B. Walczak, "Particle swarm optimization (PSO). a tutorial," *Chemometrics and Intelligent Laboratory Systems*, vol. 149, pp. 153–165, 2015, doi: 10.1016/j.chemolab.2015.08.020.

[24] D. Wang, D. Tan, and L. Liu, "Particle swarm optimization algorithm: an overview," *Soft Computing*, vol. 22, no. 2, pp. 387–408, 2018, doi: 10.1007/s00500-016-2474-6.

[25] J. Wang, S. Lu, S.-H. Wang, and Y.-D. Zhang, "A review on extreme learning machine," *Multimedia Tools and Applications*, vol. 81, no. 29, pp. 41611–41660, 2022, doi: 10.1007/s11042-021-11007-7.

[26] G.-B. Huang, D. H. Wang, and Y. Lan, "Extreme learning machines: a survey," *International Journal of Machine Learning and Cybernetics*, vol. 2, no. 2, pp. 107–122, 2011, doi: 10.1007/s13042-011-0019-y.

[27] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006, doi: 10.1016/j.neucom.2005.12.126.

[28] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual understanding of convolutional neural network- a deep learning approach," *Procedia Computer Science*, vol. 132, pp. 679–688, 2018, doi: 10.1016/j.procs.2018.05.069.

[29] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, 2018, doi: 10.1016/j.patcog.2017.10.013.

[30] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, pp. 1–28, 2020, doi: 10.1016/j.physd.2019.132306.

[31] M. Kaur and A. Mohta, "A review of deep learning with recurrent neural network," in *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, IEEE, 2019, pp. 460–465, doi: 10.1109/ICSSIT46314.2019.8987837.

[32] G. V. Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5929–5955, 2020, doi: 10.1007/s10462-020-09838-1.

[33] G. Shen, Q. Tan, H. Zhang, P. Zeng, and J. Xu, "Deep learning with gated recurrent unit networks for financial sequence predictions," *Procedia Computer Science*, vol. 131, pp. 895–903, 2018, doi: 10.1016/j.procs.2018.04.298.

[34] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: unbiased boosting with categorical features," *arXiv-Computer Science*, pp. 1–23, 2017.

[35] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121–167, 1998, doi: 10.1023/A:1009715923555.

[36] R. G. Brereton and G. R. Lloyd, "Support vector machines for classification and regression," *The Analyst*, vol. 135, no. 2, pp. 230–267, 2010, doi: 10.1039/B918972F.

[37] Facebook Open Source, "Prophet-forecasting at scale," *Facebook*. 2023. Accessed: Jan. 31, 2024. [Online]. Available: https://facebook.github.io/prophet/

[38] X. Yan and X. G. Su, *Linear regression analysis*. London, United Kingdom: World Scientific, 2009, doi: 10.1142/6986.

[39] D. Maulud and A. M. Abdulazeez, "A review on linear regression comprehensive in machine learning," *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 140–147, 2020, doi: 10.38094/jastt1457.

[40] A. Tarsitano and I. L. Amerise, "Short-term load forecasting using a two-stage sarimax model," *Energy*, vol. 133, pp. 108–114, 2017, doi: 10.1016/j.energy.2017.05.126.

[41] Y. Tsuruoka, J. Tsujii, and S. Ananiadou, "Stochastic gradient descent training for l1-regularized log-linear models with cumulative penalty," in *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1 - ACL-IJCNLP '09*, Morristown, USA: Association for Computational Linguistics, 2009, doi: 10.3115/1687878.1687946.

[42] P. Netrapalli, "Stochastic gradient descent and its variants in machine learning," *Journal of the Indian Institute of Science*, vol. 99, no. 2, pp. 201–213, 2019, doi: 10.1007/s41745-019-0098-4.

[43] C. J. Willmott, "On the evaluation of model performance in physical geography," in *Spatial Statistics and Models*, Dordrecht: Springer Netherlands, 1984, pp. 443–460, doi: 10.1007/978-94-017-3048-8_23.

[44] J. Gaboitaolelwe, A. M. Zungeru, A. Yahya, C. K. Lebekwe, D. N. Vinod, and A. O. Salau, "Machine learning based solar photovoltaic power forecasting: a review and comparison," *IEEE Access*, vol. 11, pp. 40820–40845, 2023, doi: 10.1109/ACCESS.2023.3270041.

[45] N. Bakkalbasi, K. Bauer, J. Glover, and L. Wang, "Three options for citation tracking: google scholar, scopus and web of science," *Biomedical Digital Libraries*, vol. 3, no. 1, 2006, doi: 10.1186/1742-5581-3-7.

[46] A. Nikulins *et al.*, "Deep learning for wind and solar energy forecasting in hydrogen production," *Energies*, vol. 17, no. 5, 2024, doi: 10.3390/en17051053.

[47] A. Salari, H. Shakibi, M. A. Soleimanzade, M. Sadrzadeh, and A. Hakkaki-Fard, "Application of machine learning in evaluating and optimizing the hydrogen production performance of a solar-based electrolyzer system," *Renewable Energy*, vol. 220, 2024, doi: 10.1016/j.renene.2023.119626.

[48] K. Sareen, B. K. Panigrahi, T. Shikhola, and R. Nagdeve, "Deep learning solar forecasting for green hydrogen production in india: a case study," *International Journal of Hydrogen Energy*, vol. 50, pp. 334–351, 2024, doi: 10.1016/j.ijhydene.2023.08.323.

[49] C. Emeksiz and M. Tan, "An innovative adaptive hybrid prediction model based on deep learning methods (CEEMDAN-GRU) for green hydrogen production, in case study: turkey," *International Journal of Hydrogen Energy*, vol. 52, pp. 1266–1279, 2024, doi: 10.1016/j.ijhydene.2023.11.026.

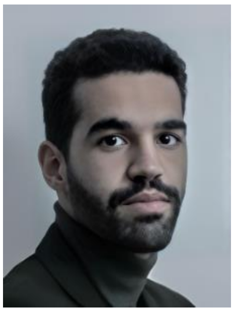
[50] A. A. Alhussan *et al.*, "Green hydrogen production ensemble forecasting based on hybrid dynamic optimization algorithm," *Frontiers in Energy Research*, vol. 11, 2023, doi: 10.3389/fenrg.2023.1221006.

[51] G. Cheng *et al.*, "Analysis and prediction of green hydrogen production potential by photovoltaic-powered water electrolysis using machine learning in china," *Energy*, vol. 284, 2023, doi: 10.1016/j.energy.2023.129302.

[52] A. Javaid *et al.*, "Forecasting hydrogen production from wind energy in a suburban environment using machine learning," *Energies*, vol. 15, no. 23, 2022, doi: 10.3390/en15238901.

[53] S. A. Haider, M. Sajid, and S. Iqbal, "Forecasting hydrogen production potential in islamabad from solar energy using water electrolysis," *International Journal of Hydrogen Energy*, vol. 46, no. 2, pp. 1671–1681, 2021, doi: 10.1016/j.ijhydene.2020.10.059.

BIOGRAPHIES OF AUTHORS



Mohamed Yassine Rhafes received an engineering degree in software engineering from the National School of Applied Sciences of Oujda at the University Mohammed Premier, Oujda, Morocco. Subsequently, then he worked as a software engineer. Currently, he is a Ph.D. student at the MATSI laboratory, Higher School of Technology (ESTO), at the University Mohammed Premier, Oujda, Morocco. His research focuses on artificial intelligence in renewable energy, particularly in the field of green hydrogen. He can be contacted at email: mohamedyassine.rhafes@ump.ac.ma.

Omar Moussaoui received his Ph.D. in Computer Science at the University of Cergy-Pontoise France in 2006. He is an Associate Professor at the Higher School of Technology (ESTO) of the University Mohammed Premier, Oujda – Morocco. He has been a member of the Computer Engineering Department of ESTO since 2013. He is currently director of the MATSI Research Laboratory. His research interests lie in the fields of IoT, AI, wireless networks, and cybersecurity. He has actively collaborated with researchers in several other computer science disciplines. He participated in several scientific & organizing committees of national and international conferences. He served as reviewer for numerous international journals. He has more than 40 publications in international journals and conferences and several co-authored book chapters, and he has h-index 13. He is an instructor for CISCO Networking Academy on CCNA Routing & Switching and CCNA Security. He can be contacted at email: o.moussaoui@ump.ac.ma.

Maria Simona Raboaca is working as a Researcher at to National Research and Development Institute for Cryogenics and Isotopic Technologies ICSI Rm. Valcea, Hydrogen and Fuel Cell Department. Her Ph.D. is "Theoretical and practical contribution regarding to sustain with hybrid energy a passive house" in Faculty of Building Services Engineering in Technical University of Cluj-Napoca, Romania. Now, she is a project manager at ICSI to project "Smart conductive charging station, fixed and mobile, for electric propulsion transportation (SMiLE-EV)" proposes the deployment of fixed and mobile EV & PHEV charging stations to meet the mobility needs of tomorrow's society and to prepare active/potential industrial partners for knowledge/technology transfer at the component or system level in prepare launching new products. She has been contributing to the field of renewable energy, green buildings, passive house concept, hydrogen energy and stationery and mobile applications. She is the author and co-author of more technical papers in scientific conference proceedings and ISI journals. She can be contacted at email: simona.raboaca@icsi.ro.