EPH–International Journal of Science & Engineering
Published by: E-Pub Solutions Private Limited
ISSN (Online): **2454-2016**
Volume 11 Issue 01 March 2025

MACHINE LEARNING-DRIVEN THERMODYNAMIC PREDICTION OF HYDROGEN STORAGE PERFORMANCE IN METAL HYDRIDES FOR SUSTAINABLE ENERGY APPLICATIONS

Dr. Abhishek Kulshreshtha[1*], Dr. Rituparna Ghosh[2]

[1]Department of Chemical Engineering, Indian Institute of Technology BHU, Varanasi, India*
School of Advanced Materials and Green Energy Technologies, Jadavpur University, Kolkata, West Bengal, India rituparna.ghosh.research@gmail.com

Article History:

Article Type: **Research**
Received Date: **20/12/2024**
Revised Date: **20/01/2025**
Accepted Date:18/**02/2025**
Published Date: **17/03/2025**

Keywords: hydrogen storage, metal hydrides, machine learning, Gradient Boosting, materials informatics

ABSTRACT

Efficient hydrogen storage remains a major challenge in the development of sustainable energy systems. Metal hydrides have emerged as promising solid-state hydrogen storage materials because of their favorable thermodynamic stability and reversible hydrogen absorption behavior. This study aimed to predict hydrogen storage performance in metal hydrides using machine learning techniques and to identify the variables most strongly influencing Hydrogen Weight Percent. A quantitative computational framework was implemented using secondary thermodynamic data containing 772 observations and 13 variables. Data preprocessing included missing-value treatment, categorical encoding, and feature scaling prior to model development. Three regression algorithms, namely Linear Regression, Random Forest, and Gradient Boosting, were evaluated using R-squared, mean absolute error, root mean square error, and five-fold cross-validation. The results showed that Gradient Boosting achieved the highest predictive performance with an R squared value of 0.9088, a mean absolute error of 0.1182, and a root mean square error of 0.2194. Random Forest also demonstrated strong predictive capability, whereas Linear Regression produced comparatively lower accuracy. Cross-validation results confirmed the robustness and generalization capability of the ensemble learning models. Feature importance analysis revealed that HtoM was the most influential predictor, contributing 57.77% of the total model importance, followed by Material Class, Temperature °C, and Pressure Atmospheres Absolute. The findings demonstrate that machine learning provides an efficient computational approach for hydrogen storage prediction and supports accelerated screening of metal hydride materials for sustainable energy applications.

1. Introduction

This reliance on fossil fuels has strengthened the awareness of environmental degradation and climate instability for industrial activities, transportation, and electricity generation. The world's energy demand is growing continuously, which has a significant impact on GHG production and has been accelerating global warming and ecological imbalance (Ghosh & Ghosh, 2020). Meanwhile, the high use of traditional energy sources has led to concerns about resource depletion, energy security and sustainability. As a result of the increased environmental impact of fossil fuel usage, global interest in cleaner and more sustainable energy options that can power future economic and industrial growth has intensified (Depren et al., 2022). Hydrogen is currently one of the most promising substitutes in the current transition to a low-carbon energy system. Hydrogen, as a high gravimetric energy density carrier, is being seen more and more as a potential energy carrier for the sustainable energy infrastructure of the future (McCay & Shafiee, 2020). It could be used in a number of sectors such as transport, industrial processing, power generation or for the storage of renewable energy. It is believed that the future use of hydrogen in the energy sector will play a major role in decarbonisation plans and carbon emission savings in energy-intensive sectors (Kyriakopoulos, 2021).

The importance of hydrogen has been further reinforced in the framework of the energy transition by recent developments in clean energy technologies. Applications that are not easily electrifiable provide opportunities for hydrogen-based systems for renewable energy integration and storage at a large scale (Johnson et al., 2025). However, as promising as it may be, there are still a number of technical and economic challenges to the realization of hydrogen technologies. Efficient hydrogen storage is one of these, and it is one of the most important hurdles for the commercialization and application of hydrogen (Hassan et al., 2024). There are several requirements that a hydrogen storage system must meet when it comes to the operation, such as high storage density, thermodynamic stability, reversibility, safety and economic feasibility. Energy losses and risks to operations are common when using traditional storage methods like compressed gas and cryogenic liquid storage. Solid-state hydrogen storage using metal hydrides has garnered significant interest due to its potential to offer safer and denser hydrogen storage at relatively moderate operating conditions (Akpasi et al., 2025). Metal hydrides have attractive thermodynamic and kinetic attributes to absorb and release H2, and they are potential materials for future energy storage applications.

As the number of new materials for hydrogen storage has become more complicated, there is a greater need for more sophisticated computational methods that can speed up the process of finding new materials while predicting their performance. The characterization process of hydride materials can be costly and laborious, as well as limited by the wide compositional scope that exists in today’s alloy systems. Due to this, materials informatics and machine learning have gained popularity as powerful tools for gaining understanding about the complicated nature of materials and successfully predicting their characteristics (Li & Zheng, 2023). Through the integration of data science approaches into materials engineering, significant insights have been obtained through large thermodynamic databases, providing an improved selection of materials for energy purposes. Machine learning approaches have been found to be quite successful in predicting various characteristics of materials, such as composite property prediction, alloy design, thermodynamics optimization, and more. Data-driven modelling methods have proven to work well for predicting complex behaviour of materials, which is hard to be characterized using the traditional methods (empirical or theoretical) (Li et al., 2022).

Machine learning has demonstrated significant promise in predicting storage capacity, optimizing reservoir performance, and enhancing thermodynamic modeling of hydrogen systems under different operating conditions in the context of hydrogen-related systems (Mao et al., 2024). In the past few years, several studies have pointed to the applicability of the use of machine learning methods in the field of hydrogen storage materials. As data-driven methods are increasingly important in the discovery of hydrogen materials, Witman et al. (2021) showed the successful application of these methods to discover high-entropy alloy hydrides with targeted thermodynamic stability. Another area that has garnered interest in ensemble-based machine learning techniques is their ability to capture nonlinear interactions and enhance prediction performance in complex materials datasets. Mishra et

al. (2022) found that the ensemble learning techniques showed great results in the prediction of the phase and were able to capture nonlinearities among the compositional parameters in high-entropy alloy systems. In the same vein, Choudhury et al. (2020) showed the effectiveness of ensemble learning for structure prediction of multi-principal element alloys and how the ensemble of multiple predictive learners can be beneficial in materials informatics tasks.

While there has been significant advancement in the application of machine learning techniques to materials engineering, there are still critical challenges to address in predicting materials for hydrogen storage. Numerous past studies have been limited to a particular class of materials, very specialised computational simulations or limited thermodynamic conditions. In some instances, the predictive significance of thermodynamic factors like hydrogen to metal ratio, equilibrium pressure, entropy and temperature has not been fully explored in a unified way of prediction.

Furthermore, the performance of different ensemble regression algorithms has not been studied much for predicting hydrogen storage capacity using the experimentally obtained thermodynamic data. This study overcomes these limitations by creating machine learning models to predict the hydrogen storage performance of metal hydrides based on thermodynamic and material properties variables. The study then compares the ability of Linear Regression, Random Forest, and Gradient Boosting to predict hydrogen storage properties and also identifies the parameters that have the highest impact on hydrogen storage properties. The study also presents materials informatics and ensemble learning, which could be used to develop sustainable hydrogen storage technologies in the future.

2. Methodology

2.1 Research Design

In the present study, the quantitative computational research design has been applied to investigate the application of machine learning methods in estimating the capacity for hydrogen storage. The thermodynamic and material properties associated with metal hydrides have been the central focus of this study. A supervised machine learning approach has been applied to develop the relationship between the performance of hydrogen storage and the chosen set of predictor variables. This study is based on the predictive analysis technique, wherein the regression models have been developed using secondary data.

2.2 Data Source

This data set is an adaptation of that of Witman et al. (2024). This data set has the properties related to thermodynamics and hydrogen storage for metal hydrides, such as pressure, temperature, equilibrium, hydrogen/metal ratio, and hydrogen storage properties. These data have been consolidated for applications in machine learning for studies on hydrogen storage materials and sustainable energy studies. The original data set had 772 observations for 13 different parameters, including material descriptors and thermodynamic variables pertinent to the hydrogen absorption behavior. The target variable we used for the predictive modelling was the Hydrogen Weight Percent, which reflects the materials' hydrogen storage capacity.

2.3 Data Preparation and Preprocessing

To ensure data consistency for analysis and enhance the predictive power of the models, data were pre-processed before developing the models. The data set was de-indexed and de-referenced, as such variables and fields did not help in predictive analysis. The missing data in the numerical variables were imputed with the median values, while the missing data in the categorical variables were imputed with the mode values. This way, there was minimal loss of data, and the distributional structure of the variables was maintained. To make categorical data like Material Class and Composition Formula compatible with machine learning algorithms, each category was assigned a number with a label encoding method. Numerical predictor variables were then standardised by applying a feature scaling procedure to normalise the differences in magnitudes of variables and promote the convergence of the model.

2.4 Exploratory Data Analysis

Exploratory Data Analysis was used to investigate the statistical properties and association of the variables used in the study. The following descriptive statistics were computed for all the quantitative attributes: mean, standard deviation, minimum, maximum, and quartile values. Linear relationship between hydrogen storage capability and thermodynamic attributes was investigated with the help of Pearson's correlation coefficient.

2.5 Machine Learning Model Development

Three machine learning techniques using regression were developed: linear regression, random forest regressor, and gradient boosting regressor. Linear regression served as the base statistical technique, whereas random forest and gradient boosting techniques helped in understanding the nonlinearity between the thermodynamic and material parameters. The data set was split into a training set and a testing set with an 80:20 ratio. Model fitting was performed using the training dataset, while predictive performance was assessed on previously unknown observations using the testing dataset. Random state initialization has been applied to splitting the data to ensure that results are reproducible.

2.6 Model Evaluation

Several performance metrics were used to assess the model performance, such as R-squared, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). These metrics were chosen to evaluate the explanatory and predictive power. Five-fold cross-validation analysis was performed for all models developed to assess the stability and generalization. Then, the best-performing model was used for feature importance analysis (FIA) to identify the most important variables in the prediction of hydrogen storage.

3. Results

3.1 Descriptive Statistics of Hydrogen Storage Variables

Initial analysis of the statistical characteristics of the hydrogen storage data set was conducted to characterize the distribution and the range of variation of the thermodynamic variables that have been used for the analysis. Table 1 shows the descriptive statistics of the numerical variables that were part of the study.

Table 1. Descriptive statistics of hydrogen storage variables

Variable	Count	Mean	Standard deviation	Minimum	First quartile	Median	Third quartile	Maximum
Hydrogen Weight Percent	772	1.5794	0.7084	0.0000	1.2000	1.5083	1.8000	6.2000
Heat of Formation kJperMolH2	772	34.3138	19.0309	1.0300	23.6950	29.7400	39.0000	219.0000
Temperature oC	772	62.1937	97.9415	-78.0000	30.0000	30.0000	36.2500	800.0000
Pressure Atmospheres Absolute	772	7.2923	136.7704	0.0001	1.1000	1.2000	1.3775	3800.0000
Entropy of Formation JperMolH2perK	772	106.9711	22.5233	-5.1790	96.4093	107.8000	117.5187	260.1844
Equilibrium Pressure 25 °C	772	128.2288	1462.9289	0.0000	0.1041	1.1726	15.0877	37390.1287
LnEquilibrium Pressure 25C	772	-0.9752	6.6777	-65.4111	-2.2817	0.1507	2.7138	10.5292
HtoM	772	0.9613	0.2895	0.0000	0.8127	0.9600	1.0912	2.3540

The data showed a wide range for some thermodynamic parameters. The mean value of the hydrogen weight per cent was 1.58 wt.%, and the hydrogen storage capacity was as high as 6.20 wt.%, which means that both low-capacity and high-capacity hydride systems were included. There was also a significant spread in the temperature and pressure parameters, as these represent a range of various experimental conditions in which hydrogen storage materials were tested. The significant standard deviations for Equilibrium Pressure 25 °C and Pressure Atmospheres Absolute indicate great variation in the equilibrium behavior and thermodynamic operating environment of the hydrides studied.

3.2 Correlation Analysis of Thermodynamic Variables

A correlation analysis via Pearson correlation coefficients was conducted to explore the correlation between hydrogen storage properties and the thermodynamic parameters. It can be seen from the correlation matrix in Figure 1 that the presence of both weak and strong relationships among the thermodynamic variables.

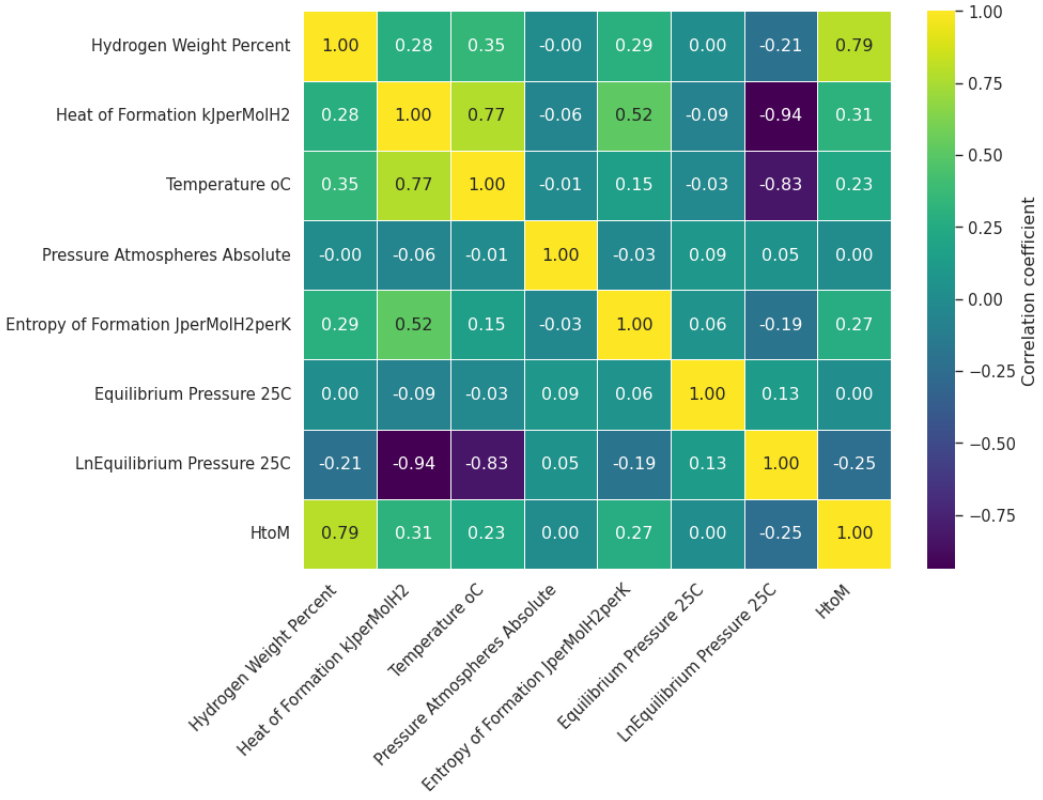

Figure 1. Correlation matrix of hydrogen storage variables

The correlation analysis revealed that the hydrogen-to-metal ratio (HtoM) was found to exhibit the highest positive correlation (r = 0.79) with the hydrogen storage capacity, which means that an increase in the HtoM significantly improves the hydrogen storage capacity. There was also a moderate positive correlation between Hydrogen Weight Per cent and Temperature ºC (r = 0.35), and between Hydrogen Weight Per cent and Entropy of Formation JperMolH2perK (r = 0.29), and between Hydrogen Weight Percent and Heat of Formation kJperMolH2 (r = 0.28). In contrast, LnEquilibrium Pressure 25 °C showed a weak negative correlation (r = -0.21) with Hydrogen Weight Per cent. The negative coefficient of correlation was found to be strong between LnEquilibrium Pressure 25°C and Heat of Formation kJperMolH2 (r = -0.94) and Temperature °C (r = -0.83), which reflects significant thermodynamic interdependence between the behaviour of both the equilibrium pressure and enthalpy and temperature conditions. It can be observed that hydrogen storage capability is dependent on complex nonlinear relations among the thermodynamic variables.

3.3 Distribution Characteristics of Hydrogen Storage Capacity

To determine the statistical characteristics of the hydrogen storage capacity in the data set, the distribution of the target variable was analyzed. The following Figure 2 is a frequency distribution of Hydrogen Weight Per cent. The distribution showed a positively skewed distribution with most results ranging from around 1 wt.% to 2 wt.%. Some materials had notably higher hydrogen storage capacities (> 3 wt.%) with a smaller proportion. The distribution deviates from normality, which suggests that there are a few more advanced hydride materials with outstanding hydrogen uptake in the high-performance region.

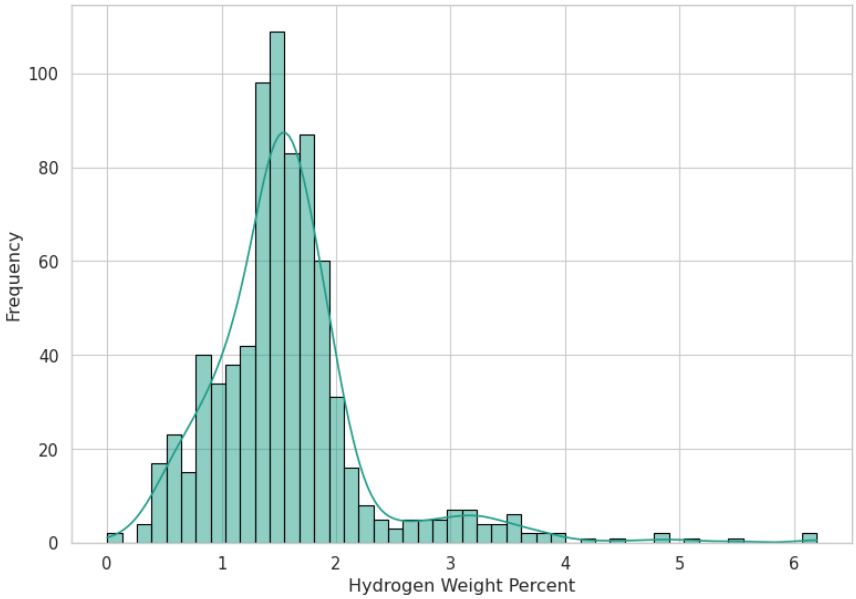

Figure 2. Distribution of Hydrogen Weight Per cent

The observed non-normality of the target variable also justifies the use of the nonlinear ensemble learning algorithms that are able to capture the complex pattern in the heterogeneous materials dataset, e.g., Gradient Boosting and Random Forest.

3.4 Machine Learning Model Performance

Three regression algorithms to calculate Hydrogen Weight Per cent based on the selected thermodynamic and material descriptors were tested. The performance of the developed machine learning models in predicting the results is summarized in Table 2.

Table 2. Performance comparison of machine learning models

Model	R^2	Mean absolute error	Root mean square error
Gradient Boosting	0.9088	0.1182	0.2194
Random Forest	0.8553	0.1266	0.2763
Linear Regression	0.5796	0.2798	0.4710

The Gradient Boosting model had the highest R-squared value of 0.9088, which means it explained around 90.88% of the Hydrogen Weight per cent variance. The model also gave the lowest value of Mean Absolute Error and Root Mean Square Error, which indicated high predictive accuracy. Random Forest was also found to have a good predictive power, but marginally inferior to Gradient Boosting. By comparison, the predictive accuracy of the Linear Regression was significantly lower, and its prediction errors were higher, indicating that linear assumptions were not sufficient to model the complex thermodynamic relationships of the phenomena of hydrogen storage behavior. The superior performance of ensemble-based learning methods suggests that nonlinear interactions between material descriptors are important to the hydrogen storage capacities.

3.5 Cross-Validation Performance of Predictive Models

Five-fold cross-validation analysis was performed for all regression models to check the model robustness and generalization capability. The average cross-validation performance results are shown in Figure 3. Gradient Boosting had the best cross-validation performance with a mean R squared value of 0.87, followed closely by Random Forest with a 0.85 mean R squared value. Linear Regression showed lower and more variable performances between validation folds.

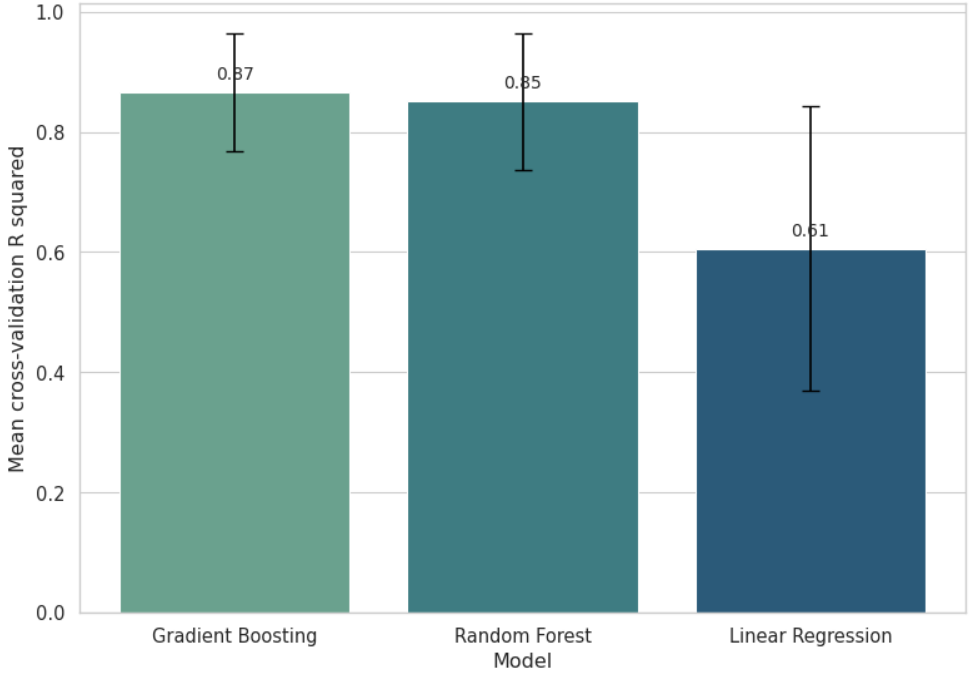

Figure 3. Five-fold cross-validation performance of machine learning models

The relatively low value of the Standard Deviation obtained for Gradient Boosting and Random Forest shows that the behavior of the models is stable and that they are not overfitting the data. The results demonstrate the validity of the ensemble learning methods to obtain reliable generalization performance when predicting hydrogen storage properties of heterogeneous thermodynamic data.

3.6 Feature Importance Analysis

The best-performing Gradient Boosting model was used to perform feature importance analysis to understand which variables were the most important towards predicting the hydrogen storage. The feature importance rankings are provided in Table 3. HtoM was the most influential predictor with an approximate contribution of 57.77% to the total model importance. Material Class, Temperature ºC and Pressure Atmospheres Absolute also played significant roles in model prediction. On the other hand, the dependencies of Equilibrium Pressure 25 °C and Entropy of Formation JperMolH2perK were relatively small in the predictive model.

Table 3. Feature importance ranking for Hydrogen Weight Percent prediction

Predictor variable	Importance score	Importance percentage
HtoM	0.5777	57.7653
Material Class	0.1346	13.4569
Temperature ºC	0.0990	9.8975
Pressure Atmospheres Absolute	0.0908	9.0824
Composition Formula	0.0422	4.2163
Heat of Formation kJperMolH2	0.0271	2.7066
LnEquilibrium Pressure 25 ºC	0.0119	1.1915
Entropy of Formation JperMolH2perK	0.0118	1.1757
Equilibrium Pressure 25 °C	0.0051	0.5077

The high value of HtoM shows how crucial it is to fill the metal lattice with hydrogen for the storage capacity. The effect of material class and thermodynamic operation parameters further highlights the role of material structure and operating environments on the hydrogen absorption behavior.

3.7 Prediction Accuracy and Residual Analysis

Gradient Boosting model predictive reliability was also tested by actual predicted and residual analyses. The plot depicting the results of the hydrogen weight percentage is shown in Figure 4 below. From the plot, there was a good correlation between the experimental and the predicted line of fit, which indicates high predictability of the results. However, a few moderate deviations in high hydrogen storage were noted, but this further supports the high predictive capability of the model for a wide variety of hydrides.

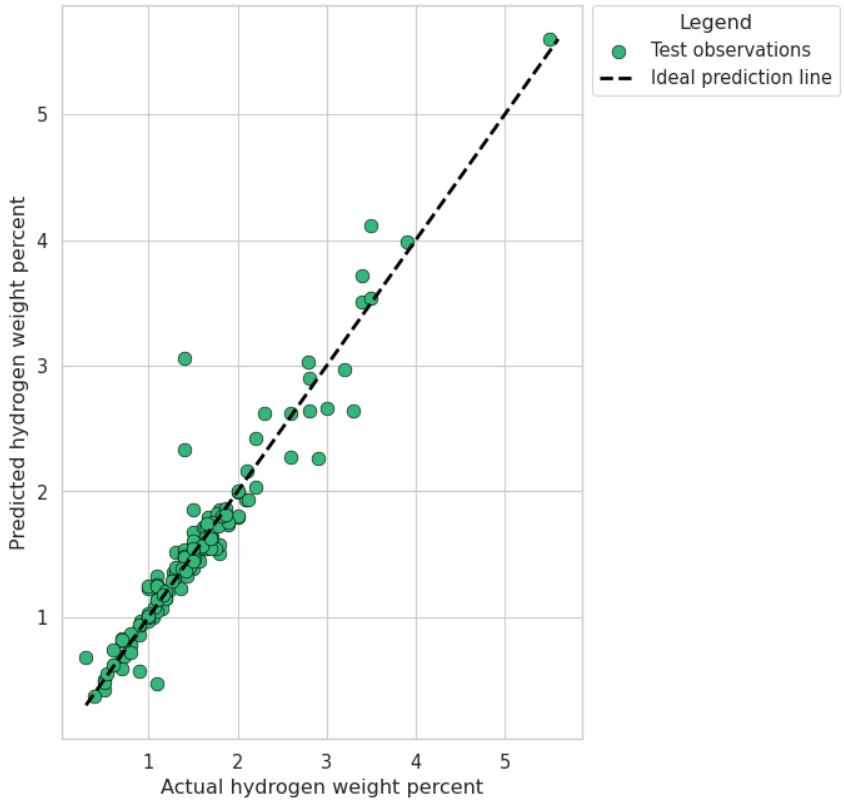

Figure 4. Actual versus predicted Hydrogen Weight Per cent values for the Gradient Boosting model

Residuals were considered in order to check the stability of the model as well as predict systematic error. The residual plot of the Gradient Boosting method is presented in Figure 5. The residuals are uniformly distributed near the value of zero without any systematic patterns, which indicates the consistency of the model and stability in predictions. The residuals are mostly located within a narrow range and prediction errors are quite low.

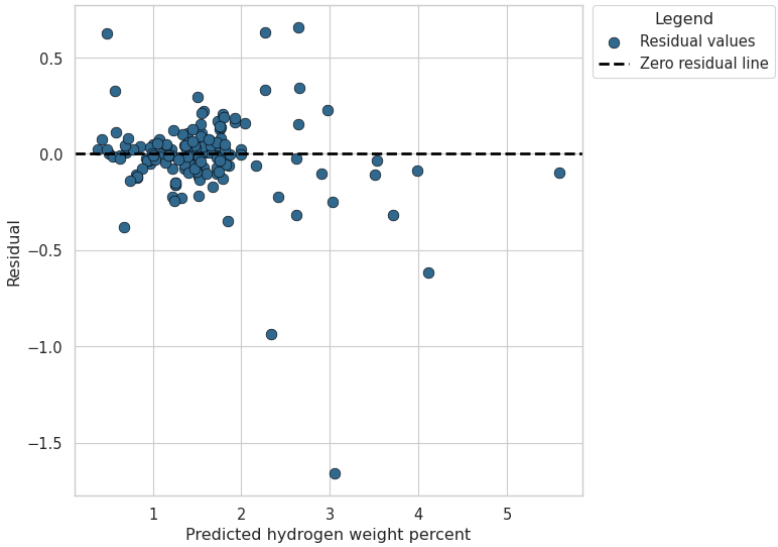

Figure 5. Residual analysis of the Gradient Boosting model

While some outliers could be seen at the high levels of predictions, there were no noticeable signs of heteroscedasticity and curvature. These findings prove the effectiveness and stability of the machine learning model that was created for hydrogen storage prediction.

4. Discussion

From the findings obtained from the analyses conducted, it can be said that both compositional and thermodynamic aspects influence the hydrogen storage capacity of a metal hydride significantly. The variation of the range of materials included in the present study was established through descriptive analysis, revealing that the range of Hydrogen Weight Per cent varied between 0.00 and 6.20 wt.%. This range can be described as an appreciable hydrogen storage capacity range, since the majority of the materials fall into the range of 1–2 wt.%, while only a few have relatively high capacity. The distribution of the materials based on their hydrogen capacity indicates the challenges involved in selecting the right hydrides that possess high gravimetric capacity. From the correlation analysis, Hydrogen Weight Per cent has been observed to be positively correlated with several features, with HtoM being the one with the best correlation. This implies that the hydrogen-to-metal ratio (H/M) is the main determinant of storage capacity, as it determines the hydrogen content in the storage structure. Moreover, temperature, Entropy of Formation and Heat of Formation were found to exhibit positive moderate relationships, indicating that thermal and thermodynamic properties play a role in the hydrogen storage process, although they do not operate alone. The negative correlation between LnEquilibrium Pressure 25 °C and Hydrogen Weight Per cent also suggests that pressure may have an influence on hydrogen storage capacity.

A linear relationship cannot provide an adequate description of the process of hydrogen storage capacity, since Machine Learning revealed more about it. From the findings, the predictive performance of Gradient Boosting was the highest, followed by Random Forest. The last in terms of accuracy was Linear Regression, meaning that understanding hydrogen storage capacity requires consideration of nonlinear interactions between Material Class, HtoM, Pressure, Temperature and other thermodynamic descriptors. The findings of cross-validation also confirm that the predictive power of the ensemble model, Gradient Boosting, was consistent throughout the validation process. The importance of the features was estimated to understand the predictive process. HtoM was revealed to be the most important feature, accounting for more than 50% of the predictive importance. Moreover, Material Class, Temperature and Pressure made significant contributions to the predictions, indicating that both structural properties and conditions during hydrogenation can influence the hydrogen storage capability of alloys. Residual and predicted versus actual plots can be used to validate the quantitative results provided by the model. The results revealed that the actual and predicted hydrogen capacities showed close agreement.

In conclusion, thermodynamic and compositional factors play a significant role in hydrogen uptake and hydrogen storage capacity, which corroborates the findings of Floriano et al. (2020). The significance of the interaction with hydrogen at the interface and at the atomic scale is also corroborated by the evidence of fast exchange of hydrogen, hydride, and proton species at the interface of transition metal oxides, which is a clear indication that the local chemical environment plays a key role in the performance of hydrogen-related materials (Wu et al., 2021). The relevance of the recorded variables: pressure and temperature is reflected in thermodynamic studies of hydrogen storage systems, where self-pressurization and thermal behavior significantly influence the stability of storage and response of the system (He et al., 2025). This excellent performance of Gradient Boosting is also in line with materials informatics studies that demonstrate good predictive capabilities of XGBoost-based models for complex functional material properties, where there exist nonlinear dependencies (Gong et al., 2022).

Additionally, recent advancements in predicting material properties, such as ensemble graph-based methods, which enhance the predictive accuracy by capturing more complex relationships between material structures and properties (Rahman et al., 2024), further underscore the value of ensemble-based learning in this work. This also holds for ferroelectric oxide systems, showing how energy landscape and defect-related descriptors can be important in understanding the behavior of materials, with the general conclusion that properties of materials are often a result of the interplay of thermodynamic and structural parameters (Azevedo Antunes et al., 2024). The present results also corroborate the findings of thermoelastic and heat-transfer studies (Abouelregal et al., 2023), where the nonlocal thermal and memory-dependent mechanisms are found to have a significant influence on the material response, thus highlighting the importance of nonlinear modeling frameworks in the study of thermodynamic material. The importance of predictive modeling in hydrogen energy materials is further emphasized by computational studies of double hydrides, which have demonstrated that the performance of hydrogen-related materials can be estimated using theoretical descriptors before the materials are even manufactured (Hayat & Khalil, 2023). Further, computational screening of metal oxides for H2 storage has proved helpful in narrowing down the number of candidate materials and prioritizing experiments (Goncalves et al., 2022).

The results have important implications for hydrogen storage material screening and sustainable energy engineering. The high accuracy of Gradient Boosting implies that machine learning can be used to help identify promising hydride materials in the early stages without resorting to the time-consuming and expensive experimental techniques. This can help in minimizing trial and error testing and enhance the efficiency of hydrogen storage research. The material design direction is also quite practical, as the HtoM is the most important feature. Materials with good hydrogen occupancy properties can be given top priority in computational screening. Meanwhile, the material, temperature and pressure contribution shows that the storage performance is not a composition-only issue, but rather a combined material and operating system issue.

Several restrictions need to be taken into account. Secondary data were used, i.e. the study relied on the consistency and completeness of the data already reported. There were some variables that needed to be imputed, which could have caused some confusion in the modeling process. Along with that, conventional regression models were used, and more advanced chemical featurization of composition formulas was not included.

Future studies could be developed to include bigger datasets, element descriptors from the composition, and predictions that have been verified in the laboratory. Hybrid modeling strategies, which involve using thermodynamic understanding and machine learning, can also enhance the predictability and interpretability of next-generation hydrogen storage materials.

5. Conclusion

Machine learning models successfully predicted hydrogen storage performance using thermodynamic and material-related descriptors. From the study, it was evident that the Gradient Boosting algorithm offered superior performance in predictive modeling with regard to the hydrogen storage capacity, since it explained the highest percentage of the variation in Hydrogen Weight Per cent, 90.88%, while

Random Forest had also performed fairly well in the task. On the other hand, the Linear Regression algorithm offered poor results due to its incapability of capturing the nonlinear relationships that govern the hydrogen storage process. HtoM had the biggest impact on the predictive capacity, with a percentage value of 57.77%. Also, there were significant contributions from other factors such as Material Class, Temperature (°C) and Pressure (Atmospheres Absolute), demonstrating that besides intrinsic material properties, the hydrogen storage capacity is influenced by the working conditions. It is evident from the results of actual and predicted data, as well as from the residuals plot, that the created framework can be considered valid in making accurate predictions. The results show that the use of machine learning techniques may support accelerated screening and optimization of sustainable hydrogen storage materials. Moving forward, composition-derived descriptors, more extensive experimental validation and hybrid thermodynamic-machine-learning models should be used to improve interpretability and real-world applicability.

References

1. Abouelregal, A. E., Marin, M., & Öchsner, A. (2023). The influence of a non-local Moore–Gibson–Thompson heat transfer model on an underlying thermoelastic material under the model of memory-dependent derivatives. *Continuum Mechanics and Thermodynamics*, *35*(2), 545-562.
2. Akpasi, S. O., Smarte Anekwe, I. M., Tetteh, E. K., Amune, U. O., Mustapha, S. I., & Kiambi, S. L. (2025). Hydrogen as a clean energy carrier: advancements, challenges, and its role in a sustainable energy future. *Clean Energy*, *9*(1), 52-88.
3. Azevedo Antunes, L., Ganser, R., Schroeder, U., Mikolajick, T., & Kersch, A. (2024). Insights Into Curie-Temperature and Phase Formation of Ferroelectric Hf1− xZrxO2 with Oxygen Defects from a Leveled Energy Landscape. *Advanced Materials Interfaces*, *11*(2), 2300710.
4. Choudhury, A., Konnur, T., Chattopadhyay, P. P., & Pal, S. (2020). Structure prediction of multi-principal element alloys using ensemble learning. *Engineering computations*, *37*(3), 1003-1022.
5. Depren, S. K., Kartal, M. T., Çelikdemir, N. Ç., & Depren, Ö. (2022). Energy consumption and environmental degradation nexus: A systematic review and meta-analysis of fossil fuel and renewable energy consumption. *Ecological Informatics*, *70*, 101747.
6. Floriano, R., Zepon, G., Edalati, K., Fontana, G. L., Mohammadi, A., Ma, Z., ... & Contieri, R. J. (2020). Hydrogen storage in TiZrNbFeNi high entropy alloys, designed by thermodynamic calculations. *International Journal of Hydrogen Energy*, *45*(58), 33759-33770.
7. Ghosh, S. K., & Ghosh, B. K. (2020). Fossil fuel consumption trend and global warming scenario: Energy overview. *Glob. J. Eng. Sci*, *5*(2), 1-6.
8. Goncalves, R. B., Snurr, R. Q., & Hupp, J. T. (2022). Computational investigation of metal oxides as candidate hydrogen storage materials. *The Journal of Physical Chemistry C*, *126*(44), 18661-18669.
9. Gong, J., Chu, S., Mehta, R. K., & McGaughey, A. J. (2022). XGBoost model for electrocaloric temperature change prediction in ceramics. *NPJ Computational Materials*, *8*(1), 140.
10. Hassan, Q., Algburi, S., Sameen, A. Z., Jaszczur, M., & Salman, H. M. (2024). Hydrogen as an energy carrier: properties, storage methods, challenges, and future implications. *Environment Systems and Decisions*, *44*(2), 327-350.
11. Hayat, M. S., & Khalil, R. A. (2023). Ab-initio exploration of unique and substantial computational properties of double hydrides Cs2CaTlH6, Cs2SrTlH6, & Cs2BaTlH6, for the computational manufacturing of hydrogen fuel cell: a DFT study. *Journal of Molecular Graphics and Modelling*, *125*, 108600.
12. He, Y., Ji, R., Ye, J., Liu, T., & Yang, Z. (2025). Non-equilibrium thermodynamic modeling and analysis of the self-pressurization mechanism in spherical liquid hydrogen storage tanks. *Applied Thermal Engineering*, 129167.
13. Johnson, N., Liebreich, M., Kammen, D. M., Ekins, P., McKenna, R., & Staffell, I. (2025). Realistic roles for hydrogen in the future energy transition. *Nature Reviews Clean Technology*, *1*(5), 351-371.

14. Kyriakopoulos, G. L. (Ed.). (2021). *Low carbon energy technologies in sustainable energy systems*. Academic Press.
15. Li, C., & Zheng, K. (2023). Methods, progresses, and opportunities of materials informatics. *InfoMat*, *5*(8), e12425.
16. Li, M., Zhang, H., Li, S., Zhu, W., & Ke, Y. (2022). Machine learning and materials informatics approaches for predicting transverse mechanical properties of unidirectional CFRP composites with microvoids. *Materials & Design*, *224*, 111340.
17. Mao, S., Chen, B., Malki, M., Chen, F., Morales, M., Ma, Z., & Mehana, M. (2024). Efficient prediction of hydrogen storage performance in depleted gas reservoirs using machine learning. *Applied Energy*, *361*, 122914.
18. McCay, M. H., & Shafiee, S. (2020). Hydrogen: An energy carrier. In *Future energy* (pp. 475-493). Elsevier.
19. Mishra, A., Kompella, L., Sanagavarapu, L. M., & Varam, S. (2022). Ensemble-based machine learning models for phase prediction in high entropy alloys. *Computational Materials Science*, *210*, 111025.
20. Rahman, C. M. A., Bhandari, G., Nasrabadi, N. M., Romero, A. H., & Gyawali, P. K. (2024). Enhancing material property prediction with ensemble deep graph convolutional networks. *Frontiers in Materials*, *11*, 1474609.
21. Witman, M., Allendorf, M., & Stavila, V. (2024). Database for machine learning of hydrogen storage materials properties (0.0.5) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10680097
22. Witman, M., Ek, G., Ling, S., Chames, J., Agarwal, S., Wong, J., ... & Stavila, V. (2021). Data-driven discovery and synthesis of high entropy alloy hydrides with targeted thermodynamic stability. *Chemistry of Materials*, *33*(11), 4067-4076.
23. Wu, S., Tseng, K. Y., Kato, R., Wu, T. S., Large, A., Peng, Y. K., ... & Tsang, S. C. E. (2021). Rapid interchangeable hydrogen, hydride, and proton species at the interface of transition metal atom on oxide surface. *Journal of the American Chemical Society*, *143*(24), 9105-9112.