A Neural Network Approach to Estimating Natural Hydrogen from Geochemical Reactions

ABSTRACT

Hydrogen has been marked as a replacement for fossil fuel in this dispensation of energy transition. In the quest to meet demands for clean energy, various methods have been utilised to produce energy. Still, recent discoveries show that natural hydrogen from water-rock interactions can be a cost-effective and viable approach to meet growing demands. Estimating natural hydrogen is complex when considering the use of thermodynamic models, thus this work explores the contribution of neural networks in the modelling and prediction of natural hydrogen concentrations. Four neural network models (SNN, BPNN, ELM, and RBFNN) were used and assessed based on seven evaluation metrics (MSE, RMSE, R, R^2, MAE, NRMSE, and NSE). Results show that the method, Extreme Learning Machine outperforms the other three on all metrics, with MSE, RMSE, R, R^2, MAE, NRMSE, and NSE values of 0.002730, 0.052246, 0.883532 0.780629, 0.045348, 0.116988 and 0.983123 respectively on the train set, and 0.000001, 0.001103, 0.999994, 0.999987,0.001047, 0.001138, and 0.999992 respectively on the test set. We posit the strength of the ELM to its simple architecture and speed in training, and its strength in aligning with this task. In all, the research proves useful in using neural networks for modelling and predicting natural hydrogen concentrations and could help researchers and policymakers in their endeavours related to clean energy and environmental sustainability.

Keywords: Spiking Neural Networks, Hydrogen production, Extreme Learning Machines, Clean Energy, Environmental Sustainability.

1.0 Introduction

In recent times, many countries have become aware of the effects of a deteriorating climate on the environment and human life. This knowledge and awareness have triggered various forms of amendments to policies regarding energy usage, and to establish commitments, agreements have been made between countries to strictly abide by certain standards at set timeframes in the quest to ensure a clean and sustainable environment [1]. These agreements have mostly revolved around the use of fossil fuels and the main idea being conveyed is endeavouring to reduce or find alternative energy sources to fossil fuel and its derivatives. To continue, researchers hold the strong belief that fossil fuel use accounts for the recent changes in rainfall patterns, and heat strokes, together with acidic rain and decreased pH in oceans. The latter has been proven to have significant impacts on marine organisms [2], and if the focus on energy transition meanders and lags, the direct impact on the world's economy would be felt through damage to capital stock and labour supply, as well as inflation on food prices during high temperatures[3].

Many forms of solutions have been proposed to deal with the long-standing use of fossil fuel which was once seen as a saviour to several economies during its early discovery. Some of these propositions require substantive investments to fully establish and maintain [4] while others in recent times are beginning to serve as auxiliary energy sources and mitigating measures. Two of such measures which are at the forefront of reduction and mitigating strategies are CO_2 capture and storage and the production and use of hydrogen. Hydrogen falls in the argument because of its characteristic nature of meeting the demanding criteria of being sustainable, clean and to some extent cost-effective, and also serves as the basis for the production of certain chemicals (ammonia) and has also been used during the production of steel[5]. Its cost is based on the method employed for its production. To this day, most of the world's hydrogen produced is using steam methane reforming, which is quite expensive and produces CO_2 emissions [6]. It is not until recently that the method of electrolysis has gained traction. Electrolysis relies on electrical current to drive a non-spontaneous chemical reaction which involves splitting water into its constituents of hydrogen and oxygen, but the limiting factor is that any excess hydrogen needs to be stored in an avenue large enough to contain the surplus, thus the storage potential of batteries and underground rock formations has been assessed [7]. While continued research is being undertaken to reduce the cost, and enhance the safety associated with underground hydrogen storage, the evolution of natural hydrogen production is making significant strides in the race for cost-effectiveness. Hydrogen is produced in natural settings through the geological processes of, but not limited to serpentinization, radiolysis, and even with the use of steel slags [8]. The naturally occurring hydrogen, often referred to as "Gold hydrogen" is useful especially for electricity production, as evidence has it that a naturally producing hydrogen reservoir serves as a source of electricity production in a town in Mali [9].

The prerequisite for any hydrogen exploration and drilling endeavour is scientific evidence which has walked through the rigour of experimentation and justifications, and several of such have been highlighted among various authors [10], [11], [12], [13], [14], [15], [16][17], [18], [19][20]. Because natural hydrogen occurrence in natural systems is time-dependent and takes a substantial amount of time for large quantities to be realised, these lab-scale experiments in their quest to mimic the time-scale processes run for months to years, making them time-consuming. The importance of lab-scale experiments cannot be emphasized enough, but it is often required to have a fair and quick idea of the hydrogen concentration (estimates) that can be produced from solid reactants, and the use of thermodynamic models has in most cases provided reliable estimates [21], [22]. These numerical models have shown to be very effective in modelling studies regarding hydrogen production, and their inner workings involve the process of minimizing the bulk chemical system's Gibbs-free energy [22]. Furthermore, besides their use in modelling and predicting natural hydrogen production, these models have also been tested in constraining the geochemical processes that influence both hydrogen production and CO_2 storage[11]. Thermodynamic models can give reasonable predictions close to experimental data even under dynamic conditions with varying water to rock ratios and

temperature conditions [21]. Though they abstract the nuance involved in solving the equilibrium constant equation, which is important, especially for complex systems, its use requires a deep understanding of the chemical reactions that dictate the hydrogen production process and goes with the need for a good grasp of the governing thermodynamic principles that regulate the geochemical reactions.

To propose a simpler approach, this work adopts Neural Networks in modelling and predicting natural hydrogen concentrations from a geochemical process of water-steel slag interactions. The justification for the use of Neural Networks is their efficacy in acting as universal approximators for any function, and predicting outputs based on inputs and learned parameters [23]. Regarding preventive and mitigative measures to CO_2 emissions, neural networks have and continue to play significant roles in prediction studies. Here, [24] trained a BPNN to assess the performance of CO_2-enhanced oil recovery and storage in residual oil zones, while [25] considered the Elman Neural Network together with other machine learning methods in modelling the interfacial tension of a hydrogen-brine system, and even flash scheme calculations have been undertaken with the use of thermodynamic informed neural networks [26].

One striking observation in the use of deep learning in the field of natural H_2 production and CO_2 storage is the over-reliance on the Multi-Layer Perceptron Backpropagation Neural Network. Multilayer Perceptron Backpropagation Neural Networks (MLP-BPNN) are interconnected layers of nodes trained with backpropagation, using techniques like gradient descent, activation functions, weight initialization, and regularization to learn complex patterns in data, with considerations for overfitting, hyperparameter tuning, and model evaluation, applicable to regression, classification, and time-series tasks, often enhanced with ensemble methods and batch normalization, demanding sufficient data, careful architecture design, and computational resources, while addressing challenges like exploding/vanishing gradients and benefiting from techniques such as early stopping, transfer learning, and regularization. Though the method has provided good results in some instances, it is not computationally efficient, and could probably provide weights which are not the global best i.e falling into local minima or suboptimal solutions during the training process.

On the contrary, the Extreme Learning Machine allows for a non-iterative approach to finding its optimal weights and giving predictions. The ELM's architecture is similar to that of the RBFNN, in that, both are made of a single hidden layer feedforward network, though other researchers have explored deep architectures in the ELM [27]. The simple architecture allows for efficient and faster computation offering advantages in computational efficiency and simplicity compared to traditional gradient-based methods while the RBFNN also has the capability of handling complex data with many dimensions, offers faster training and testing durations, and the potential to approximate any continuous function with unlimited precision [28]. The spiking neural network also comes in with its characteristics of closely mimicking the human brain more closely than the traditional ANN and adding time dependency to its functionality while allowing for sparsity in neuronal firing.

To this end, though these deep learning methods have been applied in various jurisdictions in geoscience, they are yet to be appraised in modelling and predicting natural hydrogen from geochemical reactions. The exemplary results provided by these techniques in both CO_2 and hydrogen storage make them good candidates for their application in the field of natural hydrogen production based on geochemical reactions. Thus, their use in predicting hydrogen concentration can provide useful knowledge for researchers and policymakers. In this regard, this study seeks to fulfil two overarching objectives of

- Modelling and predicting natural hydrogen concentration from geochemical reactions of water and steel slags using neural networks
- Assessing and making comparisons of the networks based on their predictive performance.

2. Relevant studies

We collated existing literature related to the use of machine learning and deep learning techniques in predicting hydrogen produced from natural processes and the use of the techniques in hydrogen storage. The techniques employed could come in handy when exploring other ventures related to natural hydrogen exploration. Typically, the use of deep learning for modelling geochemical processes was lacking, thus the literature considered general hydrogen production from natural processes including gasification, and biomass pyrolysis. Also, we explored recent works related to the use of deep learning in hydrogen storage.

3. Data and methods used.

3.1 Data Used

Datasets obtained from experimental work undertaken by [8] were used to build and test the models. The experiment undertaken by [8] though describing the production of hydrogen from steel slags, is a classical representation of geochemical reactions occurring in nature to produce hydrogen and thus fits the purpose of this work. The dataset includes reactant mass (mg), water (mg), temperature (K), duration (h) and magnetite weight (wt %) as predictors, and hydrogen concentration (mol/kg) as the dependent variable. Table 1 describes the statistics of the dataset.

Table 1. Descriptive Statistics of the entire Dataset used in training and testing of the models (n = 31)

Statistic	Predictors					Output Variable
	M reactant (mg)	M water (mg)	Temperature (K)	Duration (h)	Magnetite (wt.%)	H_2 (mol/kg)
Min	40.93	14.98	43.2	43.2	4.02	0.001
Max	102.6	231.3	1362	1362	13.29	0.484
Standard Deviation	13.62	35.63	66.28	365.6	2.33	0.11
Mean	78.51	80.22	554.67	288.51	8.34	0.194

Spearman's correlation analysis was performed to explore the relationship between the input variables (reactant mass (mg), water (mg), temperature (K), duration (h) and magnetite weight (wt %)) and the prediction target (hydrogen concentration).

$$SCC = \frac{\sum(x_i - \mu_x)(y_i - \mu_y)}{(n-1)\sigma_x\sigma_y}$$

- SCC represents the Spearman correlation coefficient.
- xi and yi are the individual data points.
- μx and μy are the means of xx and yy respectively.
- $\sigma x$and σy are the standard deviations of x and y respectively.
- n is the number of data points.

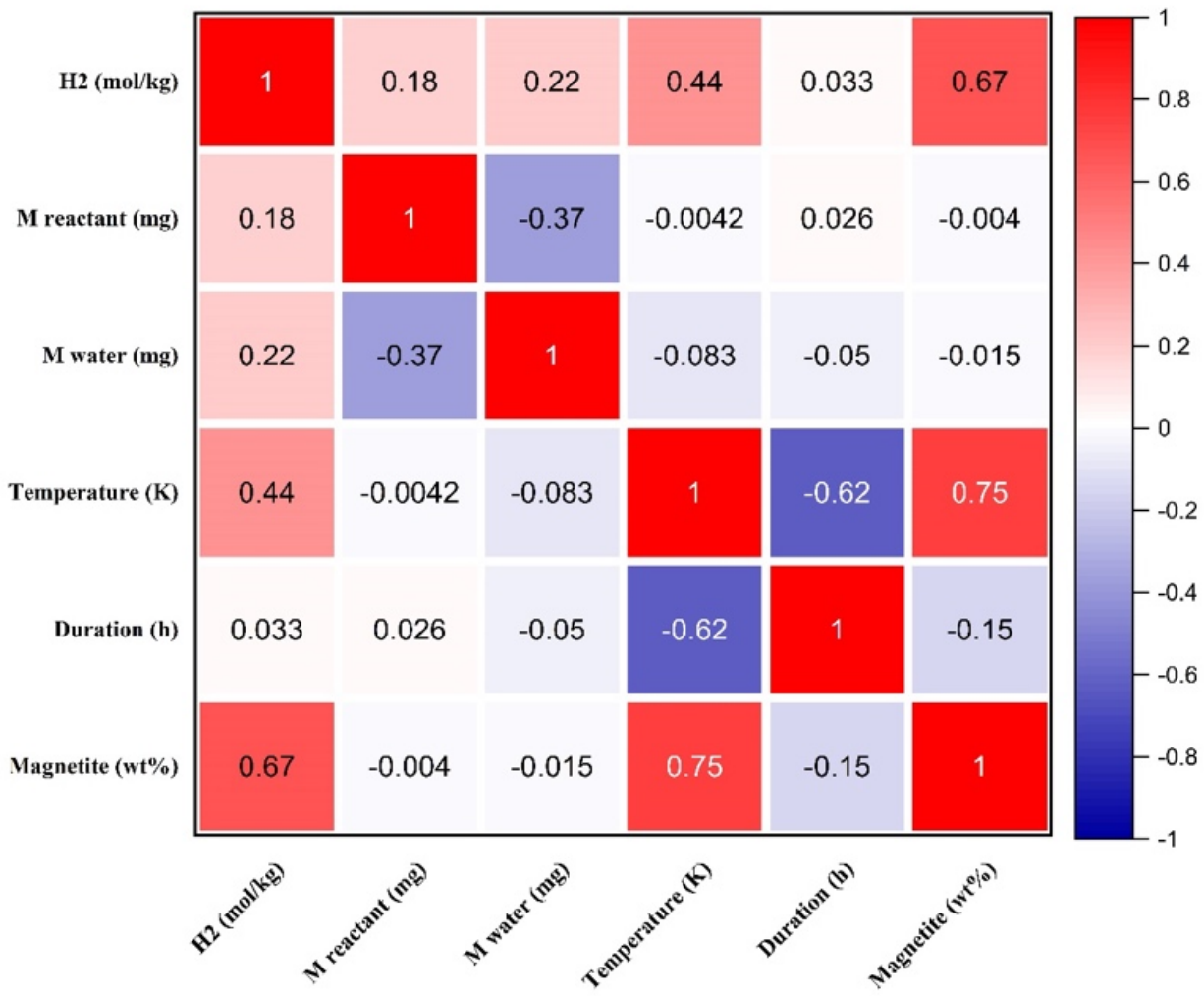

Figure 1. Spearman Correlation between variables

Table 1. Relevant Works from Literature

Objective	Model(s)	Performance	Year	Ref
Using ANN to predict dark fermentative hydrogen production from organic wastes using acidogenic mixed consortia	Response Surface Methodology (RSM), ANN and Support Vector Machine (SVM).	SVM outperformed ANN and RSM	2020	[29]
Assisted phase equilibrium analysis for producing natural hydrogen using deep learning	Thermodynamics-informed neural network (TINN)	Implementing the TINN architecture can speed up flash scheme computations by almost twentyfold.	2024	[26]
Gasification process prediction using ANN	BPNN	Pearson correlation of 0.9310 obtained with 18 hidden neurons	2022	[30]
Integrated Biomass Gasification Power Plant process parameter prediction	BPNN	The ideal hidden layer size was found to be 60 neurons, yielding the best test performance with an MSE score of 1,497 and a test R of 0.9976.	2022	[31]
Predicting hydrogen production from confectionery wastewater	BPNN with configuration of 4-12-4-2	correlation coefficient of 0.996 and average percentage error (APE) of 0.0004 on test data	2019	[32]
Biohydrogen production from starch wastewater	BPNN with configuration of (3-8-4-1)	R2-value of 0.945 for training, 0.652 for validation, and 0.791 for testing	2013	[33]
Prediction of hydrogen production and COD removal in an UASB bioreactor	Response Surface Methodology (RSM) and ANN	ANN and RSM achieved hydrogen yield of 2.22% and 9.64%, and COD removal of 1.01% and 6.34%, respectively.	2017	[34]
Enhancing biohydrogen production by Enterobacter species through artificial neural network and response surface methodology optimization.	Response Surface Methodology (RSM) and ANN	ANN outperformed RSM	2013	[35]
Modelling of hydrogen production using Nickel Loaded Zeolite	Response Surface Methodology (RSM) and ANN	The artificial neural network (ANN) displayed notable superiority, with its R-squared value nearly reaching 1, in contrast to the RSM model.	2015	[36]
Modelling Dark Fermentation of Coffee Mucilage Wastes for Hydrogen Production	Artificial Neural Network Model and Fuzzy Logic Model	Fuzzy logic outperformed ANN with R2 values of (> 0.7866) and fuzzy logic model (> 0.8485)	2020	[37]
Forecasting the retrieval of underground hydrogen storage through pore network modelling and machine learning techniques.	Least Square Fitting and Support Vector Machine (SVM)	SVM sorted rocks by high (>50%) and low (<50%) hydrogen trapping rates. Machine learning revealed that rocks with small pore-to-throat ratios and high connectivity had lower hydrogen trapping rates.	2024	[38]

4 Methods

4.1 Data Pre-processing

Before model training, the dataset was normalised using the Min-Max normalisation which scales the predictors between 0 and 1. The technique allows for early convergence and solves the issue of vanishing or exploding gradients [28] It is mathematically expressed as:

$$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} * (\max - \min) + \min \quad (3)$$

where X_{scaled} is the scaled value of *xi*, Xmin is the minimum, and Xmax is the maximum value.

Table 2. Sample of the Dataset used for training and testing the model

M reactant (mg)	M water (mg)	Temperature (K)	Duration (h)	Magnetite (wt%)	H_2 (mol/kg)
47.9	119.6	523	70.4	7.06	0.22
41.6	231.3	523	142.8	7.69	0.174
81.9	80.6	523	1362	12.5	0.484
81.4	80.5	523	329.3	9.01	0.229
81.6	80.3	623	65.3	10.9	0.254
85.9	80.6	473	497	5.59	0.214
94	80.4	673	43.2	9.77	0.323
83.6	80.7	523	785.7	9.54	0.323
81.1	80.3	573	66	7.59	0.215
81.1	81	473	69.2	5.49	0.105
80.6	80.4	673	68.3	10.1	0.265
81.33	80.62	673	68.3	8.17	0.124

4.2 Spiking Neural Networks

The proposed method (Spiking Neural Networks) is described. The benchmark methods have been highlighted extensively in various fields in the literature [39], [40], [41]. Spiking neural networks (SNNs) are a type of artificial neural network (ANN) designed to closely resemble natural neural networks. Unlike traditional ANNs, SNNs not only consider neuronal and synaptic states but also integrate the element of time into their operational framework. The concept revolves around the notion that neurons within the SNN refrain from transmitting data during each propagation cycle, a departure from typical multi-layer perceptron networks. Instead, information transmission occurs solely when a neuron's membrane potential, representing its membrane electrical charge, attains a predefined threshold value. Upon reaching this threshold, the neuron initiates firing, producing a signal that propagates to other neurons. These neurons then adjust their potentials, either increasing or decreasing, in response to this signal. A neuron model characterised by firing at the point of threshold crossing is termed a spiking neuron model.

The Leaky Integrate and Fire model is used as activation functions in the model. The leaky integrate-and-fire model stands out as the most notable among spiking neuron models. Within this model, the transient activation level, often represented as a differential equation, typically defines the neuron's state. Incoming spikes exert influence on this value, either elevating or diminishing it until the state either dissipates over time or, upon reaching the firing threshold, prompts the neuron to fire. Following the firing, the state variable undergoes a reset to a lower value. The LIF model can be mathematically represented as:

$$\tau m\ (dV/dt) = -(\ V(t) - Vrest) + Rm\ *\ I(t) \quad (1)$$

- $V(t)$ is the membrane potential of the neuron at time,t.
- τm is the membrane time constant, which represents the time it takes for the membrane potential to reach
- approximately 63.2% of its final value in response to a step input
- $Vrest$ is the resting potential of the neuron, representing the membrane potential when the neuron is not firing or receiving any input.
- Rm is the membrane resistance, which determines how easily the neuron's membrane potential changes in response to input currents.
- $I(t)$ is the input current to the neuron at time t.

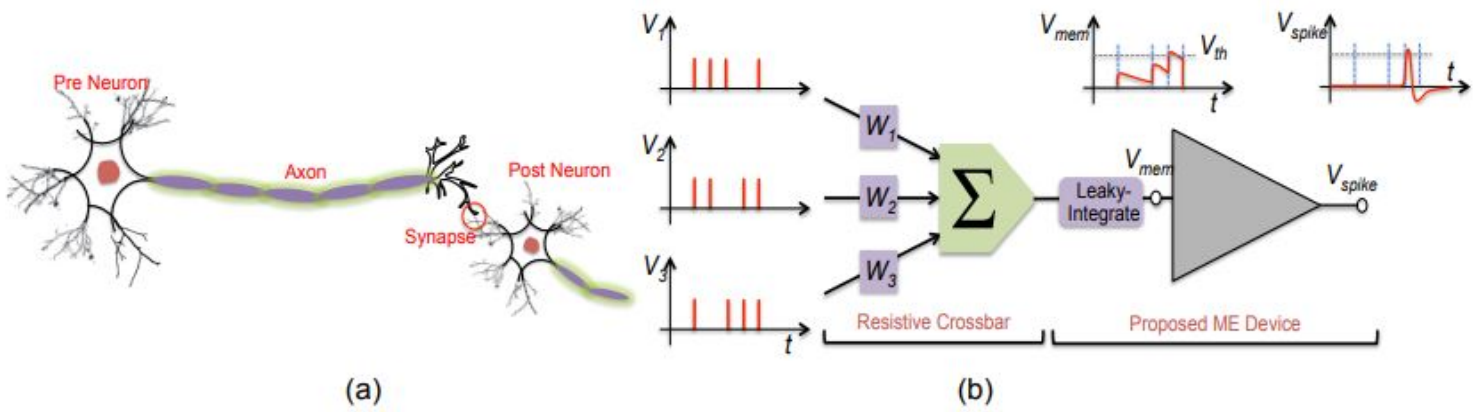

Figure 1 A representation of the Leaky Integrate and Fire Model. A) A biological neuron undergoing action potential and subsequent communication with an adjacent neuron at their synapse. B) Multiple diagrams describing the mechanism of

membrane potential accumulation and subsequent spikes as the potential reaches a threshold which is followed by a reset mechanism. Taken from [42]

In its natural state, the LIF cannot be trained using the traditional gradient descent algorithm due to the discontinuities that exist in segments of the model thus rendering them non-differential. The solution is the use of a surrogate gradient descent which is a sigmoidal function which models and updates the parameters in the LIF membrane potential. The surrogate gradient descent can be represented as:

$$\theta_{t+1} = \theta_t - \eta \nabla \theta_t L(y', y) \quad (2)$$

- θt is the parameter vector at time t.
- η is the learning rate.
- $L(y', y)$ is the loss function, measuring the discrepancy between the predicted output y' and the true target output y.
- $\nabla \theta t L(y', y)$ is the gradient of the loss function with respect to the parameter vector θt.

In surrogate gradient descent, the gradient $\nabla \theta t L(y', y)$ is computed using the surrogate function instead of the original non-differentiable activation function. This allows for the use of gradient-based optimization algorithms, such as stochastic gradient descent or Adam, to update the model parameters and minimize the loss function.

The model was trained and tested after data pre-processing, and the training process involved partitioning the data in a ratio of 70% to 30% for training and testing respectively, and subsequently trained using the batch gradient descent. The model was designed based on the features of the data, with 5 input neurons representing the predictors (reactant, water, temperature, duration, and magnetite) and an output neuron representing hydrogen concentration, with two hidden layers possessing 10 and 15 neurons respectively with embedded leaky integrate and fire activation functions. As mentioned, though the dataset represents experiments undertaken for high-purity hydrogen production using steel slags, the geochemical process involved generalises what occurs in natural settings during serpentinization reactions, as such, the developed model is suited for predicting hydrogen concentration in these natural settings. The addition of magnetite as a predictor also reflects the role of spinels in natural hydrogen production. Spinels have been proven to enhance hydrogen production in various systems [17], [29], and their inclusion makes the model a suitable representative. In the testing phase, the parameters learned during the training process (the trained model) were subsequently used on the test data to assess the model's performance on unseen data. Also, to avoid any instance of overfitting and to ensure good generalisation on the test set, the model was monitored during each epoch using the MSE as the criterion and the best and optimal MSE chosen after training.

4.3 Back Propagation Neural Network

The BPNN was considered in modelling and predicting natural hydrogen concentration. In this study, the developed model consisted of a single hidden layer with 5 neurons. The foundation of neural network training is backpropagation. It is a technique for adjusting a neural network's weights based on the error rate recorded in the previous epoch, and for batch training, the term iteration is

befitting. By properly tweaking the weights, the error rates are likely to be lowered and improve the model's reliability by broadening its applicability [43]. The weighted sum of the input variables was fed through and served as the input for the hidden layer.

In the hidden layer, the nonlinear log sigmoid transfer function was used. The transfer function is a mathematical expression that establishes a relationship between the input and output variables. This relationship can be expressed in terms of both spatial and temporal frequencies [[44]]. The spatial and temporal frequencies mentioned can be likened to wave phenomena where the former deals with the complete periods of passes undertaken by signals within a specified distance, likened to the period the input variables move through the transfer function, while the temporal function in wave-like terms refers to the time taken for passes of a wave which can also be likened to the time taken for input variables to be computed through the transfer function.

The log sigmoid function squashes the output from the input layer between a value range of 0 and 1 using the mathematical expression;

$$f(x) = \frac{1}{1+e^{-x}} \quad (3)$$

The weight optimization was executed using the Adam optimizer. The algorithm is easy to put into practice, operates efficiently in terms of computational resources, requires minimal memory, and remains unchanged even when the gradients are rescaled diagonally. Additionally, it's suitable for dealing with objectives that change over time and challenges where gradients are noisy or sparse [45].The optimizer works with the gradient descent technique, the objective is to minimise the cost function in the steepest direction by finding the partial derivative of the cost function respective to individual weights in the neural network.

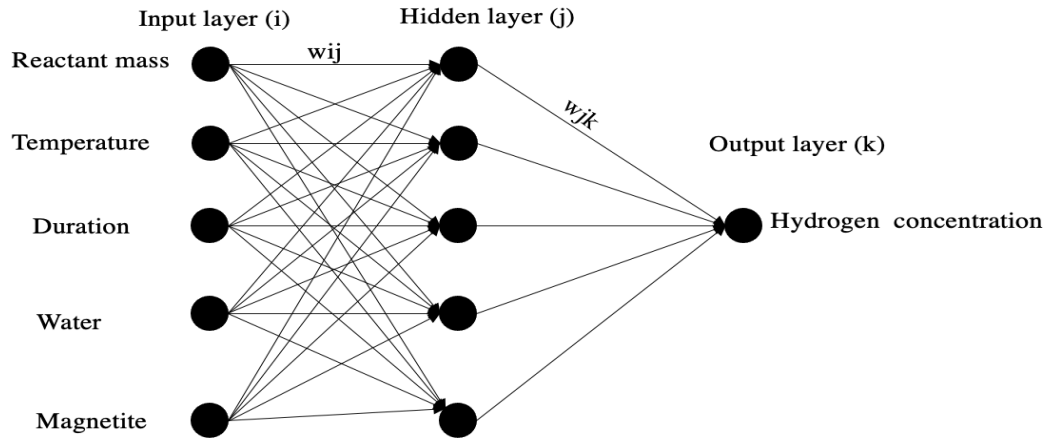

Figure 2. The BPNN architecture. All hidden layers have embedded sigmoid activation functions.

4.4 Radial Basis Function Neural Network

The RBFNN was used in modelling and predicting natural hydrogen concentration. The RBFNN is used in function approximation, classification, and prediction. In this RBFNN network, the radial basis function is used as the transfer function. Numerous RBF-based techniques have recently attracted attention in the literature, and their key benefits are the spectrum convergence rates that may be achieved utilising basis functions which are infinitely smooth in nature, geometrical adaptability, and simplicity of their implementation. In many dataset modelling applications, RBFs have performed better, especially when the dimension (D) of the space in which the dataset is located is greater or equal to 2. Studies have also demonstrated their unique benefits when the dataset is dispersed as opposed to gridded. The approach is consequently a popular option in many different fields, ranging from statistics to the approximation of partial differential equations, because of its high approximation qualities and ease of implementation. A number of radial basis functions have been used in the literature. Some of which include the; thin plate spline, gaussian, multi quadric, inverse multi quadric, and bi harmonic. In this study, the gaussian function was used in the hidden layer which consisted of 10 neurons. The RBFNN consists of three layers. The first layer receives input as linear, and the next layer is the nonlinear hidden layer. The Gaussian outputs are linearly combined in the third layer. The Gaussian activation function is characteristically shaped like a bell and is expressed in mathematical terms and its parametric form as

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- where, μ represents the centre or mean of the Gaussian distribution. It determines the location of the peak or centre of the bell curve.

- σ represents the standard deviation of the distribution. It controls the spread or width of the bell curve. Larger values of σ result in wider curves, while smaller values result in narrower curves. The width and spread were set as learnable parameters and optimised in the same manner as the weights using the Adam optimizer.

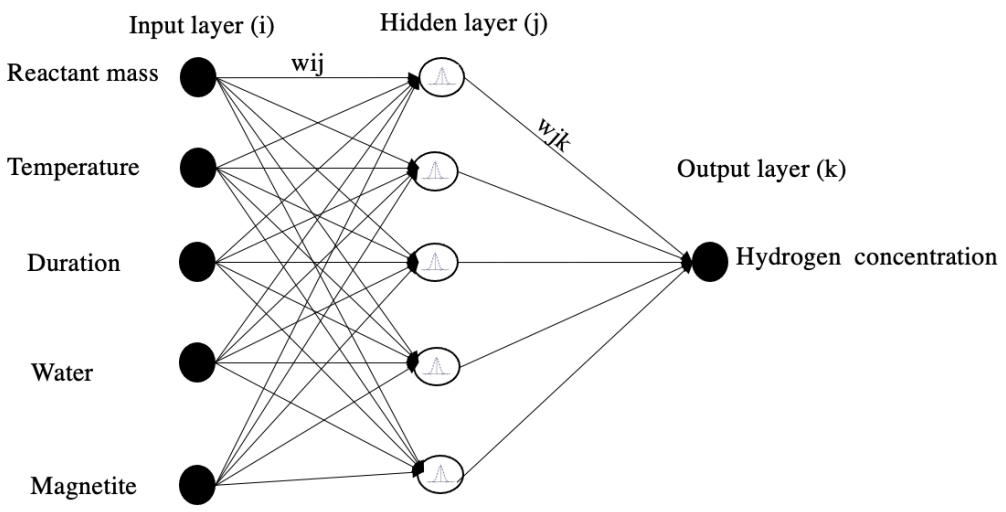

Figure 3. The Radial Basis Function Neural Network Architecture

4.5 The Extreme Learning Machine

Extreme Learning Machine (ELM) is a type of machine learning algorithm for supervised learning tasks, particularly in the context of neural networks. It was introduced by [46]. ELM is primarily used for classification and regression tasks. ELM consists of a single hidden layer, unlike traditional neural networks which may have multiple hidden layers. This simplicity helps in reducing training time and computational complexity. In ELM, the weights connecting the input layer to the hidden layer are randomly generated and fixed. These random weights are typically generated from a uniform distribution. Instead of iteratively updating the weights through backpropagation as in traditional neural networks, ELM calculates the output weights directly using a simple linear regression approach by adopting the Moore-Penrose generalised pseudo inverse to analytically determine the output weights, and this allows for fast training. The activation function used in this work is the sigmoid activation function in 9 hidden neurons. The

number of hidden neurons was selected based on the optimal results provided by the set criterion. The generalised flow of input-output in the ELM can be represented as;

$$H = X \cdot W + b \tag{4}$$

$$A = \frac{1}{1+e^{-H}} \tag{5}$$

$$\beta = (A^T A)^{-1} A^T Y \tag{6}$$

$$\hat{Y} = A \cdot \beta \tag{7}$$

5 Model Evaluation

The model performance was evaluated using 7 statistical metrics. These metrics allowed for the comparison of the effectiveness of each model. They include the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Pearson correlation (r), Correlation of determination(r^2), Mean Absolute Error (MAE), Normalized Mean Squared Error (NRMSE), and the Nash–Sutcliffe model efficiency coefficient (NSE). The mathematical representations of the evaluation metrics are described in equations (7 - 14)

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{8}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{9}$$

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \tag{10}$$

$$r^2 = \left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}\right)^2 \tag{11}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{12}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(O_i-S_i)^2}{\sum_{i=1}^{n}(O_i-\bar{O})^2} \tag{13}$$

$$\text{NRMSE} = \frac{\text{RMSE}}{\text{Max}(y)-\text{Min}(y)} \times 100\% \tag{14}$$

6 Results and Discussion

Table 3

Evaluation metrics of the models on the training data

Methods	Evaluation Metrics						
	MSE (mol/kg)	RMSE (mol/kg)	R	R^2	MAE (mol/kg)	NRMSE	NSE
BPNN	0.008074	0.089858	0.774006	0.599086	0.073380	0.200839	0.950076
SNN	0.004114	0.064140	0.817885	0.668936	0.05448028	0.143784	0.974564
RBFNN	0.008252	0.090840	0.578932	0.335162	0.076684	0.203669	0.948979
ELM	**0.002730**	**0.052246**	**0.883532**	**0.780629**	**0.045348**	**0.116988**	**0.983123**

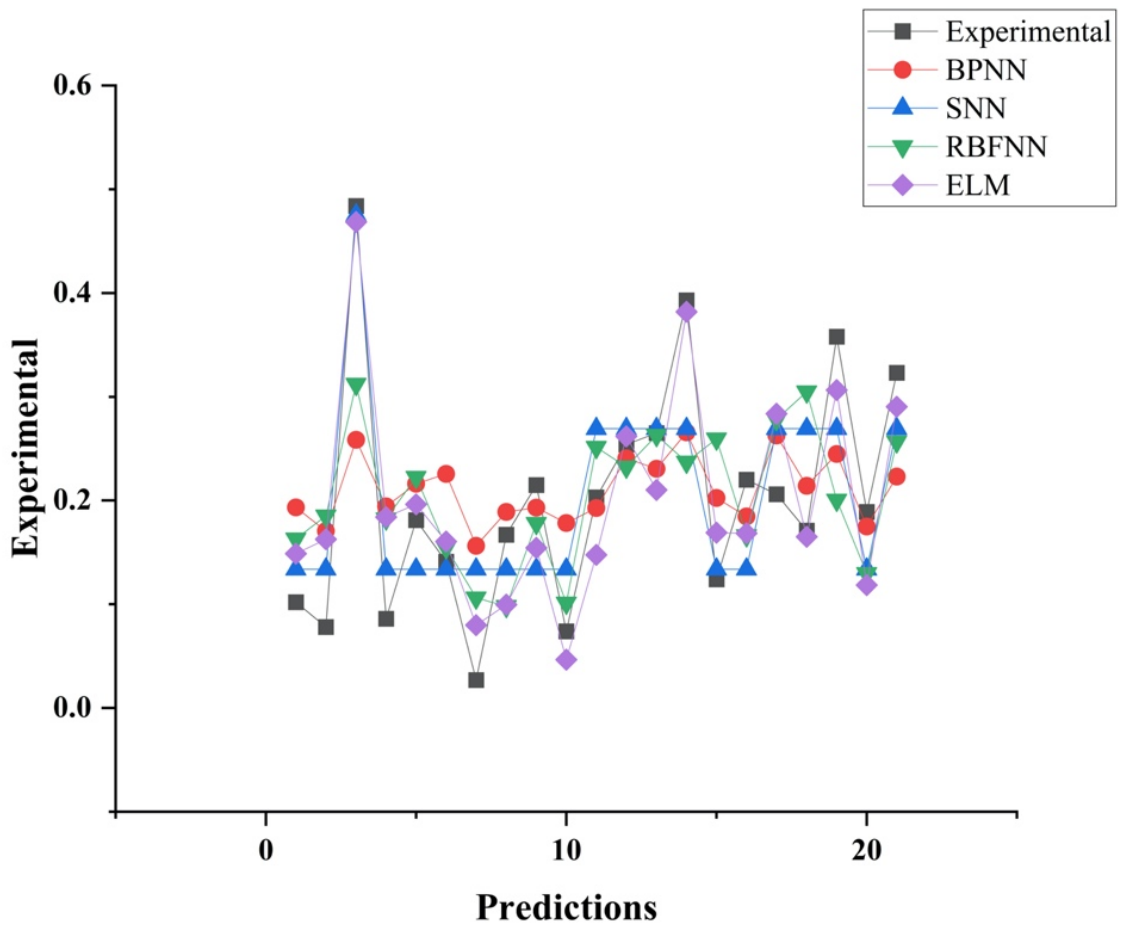

Figure 4. Plot of Experimental and Predictions on the training set

Figure 4 highlights the representation of relationships between the predictions and experimental data. The extreme learning machine successfully captures the underlying patterns better than the other models. It correctly accounts for how changes in the variables influence natural hydrogen production. The model is robust to variations and the inherent dynamics in the data, as it differentiates between random fluctuations and draws meaningful patterns, ensuring that its predictions remain accurate even in the presence of variability. Contrary to this, the BPNN exhibits a different dynamic on the training data. There could be several reasons for its poor performance and a salient one could be that the model experienced underfitting. Though underfitting could be resolved by making simpler models quite complex to improve learning, for this study, making the model quite complex in the slightest manner caused a rapid evolution of overfitting, thus a balance needed to be struck to ensure good generalisation based on the optimum architecture which provided optimal MSE value on the test set.

Table 4

Evaluation metrics of models on the test data

Methods	Evaluation Metrics						
	MSE (mol/kg)	RMSE (mol/kg)	R	R^2	MAE (mol/kg)	NRMSE	NSE
BPNN	0.006722	0.081988	0.710000	0.504100	0.067616	0.222065	0.953087
SNN	0.004028	0.063465	0.805410	0.648685	0.056543	0.173017	0.971890
RBFNN	0.003145	0.056079	0.849172	0.721094	0.050182	0.155861	0.978052
ELM	**0.000001**	**0.001103**	**0.999994**	**0.999987**	**0.001047**	**0.001138**	**0.999992**

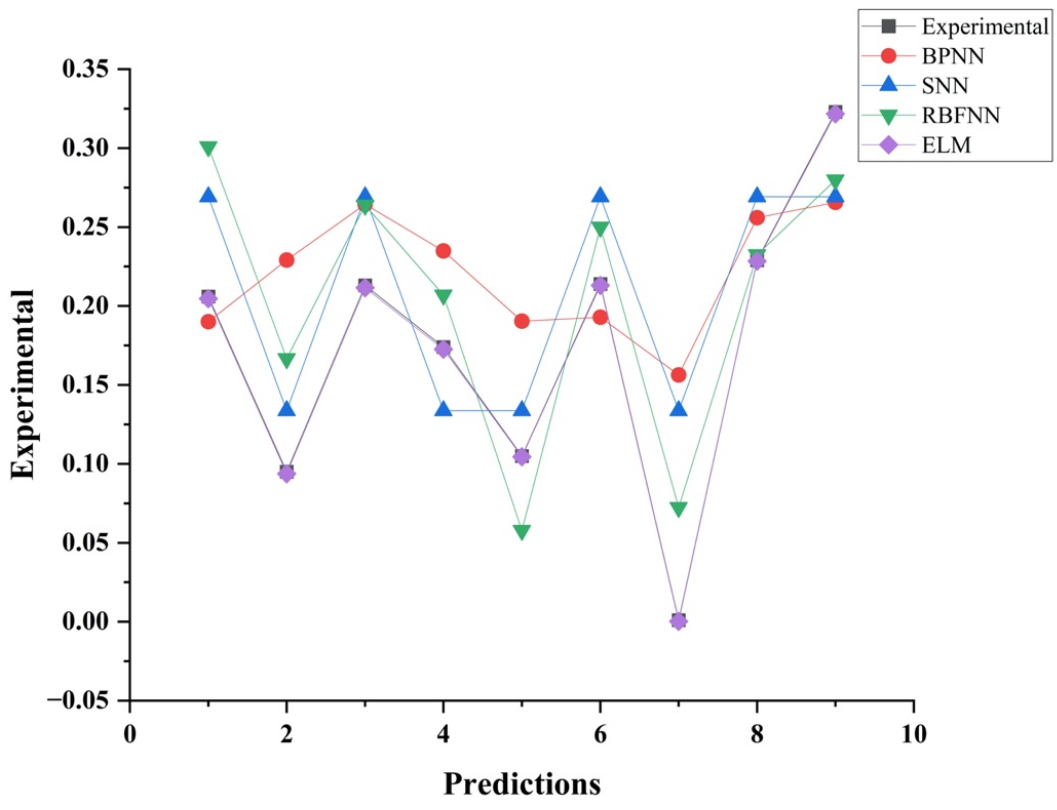

Figure 5. Plot of experimental and model predictions on the test set

The extreme learning machine also outperforms the other models on the test set just as it did on the train data. The incredible performance of the model can be attributed not only to its ability to converge faster than the other models but the relative simplicity of the algorithm in obtaining the optimal weights between the hidden and output layers. It turns out that one striking similarity that runs through the other three models is the use of the gradient descent technique to compute the optimal weights of their respective networks. This technique is prone to falling into local minima such that the so-called optimum weights are not the global optimum that could be obtained [47]. Moreso, RBFNN and BPNN often require careful tuning of parameters such as the number of hidden units, learning rate, and regularization parameters. If these parameters are not chosen optimally, the performance of these models can degrade. In contrast, ELM is less sensitive to parameter tuning, making it easier to use and yielding more consistent results across different datasets. In this study, ELM was better suited for the specific problem of natural hydrogen modelling and prediction of natural hydrogen concentration compared to RBFNN, BPNN, and Spiking Neural Networks. The inherent characteristics of ELM, such as its fast learning speed and simplicity, aligned well with the requirements of the hydrogen prediction task, leading to better performance.

Several researchers have also highlighted the robustness and effectiveness together with the speed in the learning of the extreme learning machine. [48] compared the performance of Extreme Learning Machines (ELM) and Radial Basis Function Neural Networks (RBFNN) in the domain of channel equalization. The study revealed that methods based on ELM achieved superior outcomes in terms of symbol error rate and learning speed. This outcome signified that ELM outperformed RBFNN in the context of channel equalization application. There has even been a hybrid approach of the ELM and Particle Swarm Optimization which was used by [49], in predicting the pressure of coal slurry transportation pipelines and their results show the superiority of the ELM-PSO over support

vector machine-PSO. These findings reveal the efficacy and efficiency of ELM in regression problems and have proven how simplicity can be key in some modelling problems as compared to complexity.

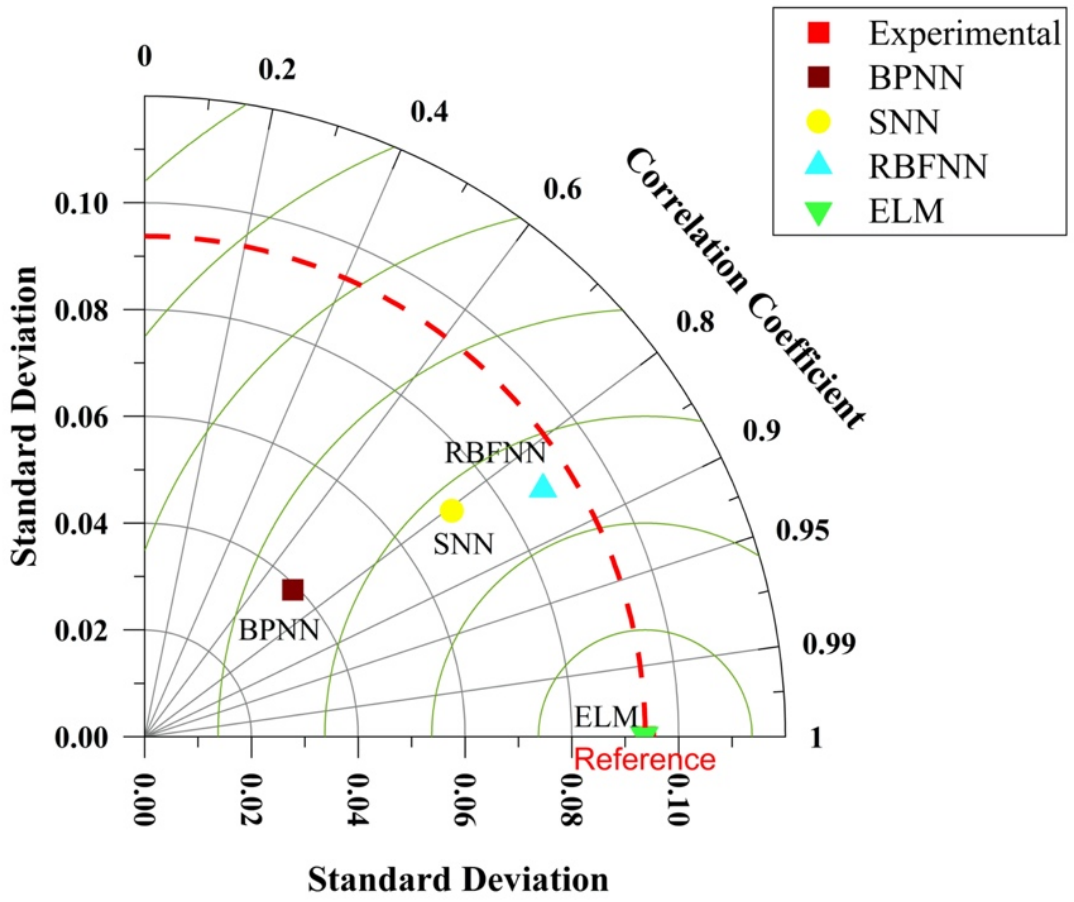

Figure 6. Taylor's diagram highlighting the comparison of the models against the experimental data on the test set. The red dashed lines show the standard deviation for the experimental data with concentric circles highlighting the central root mean squared errors, with each circle having measured values of 0.02 mol/kg

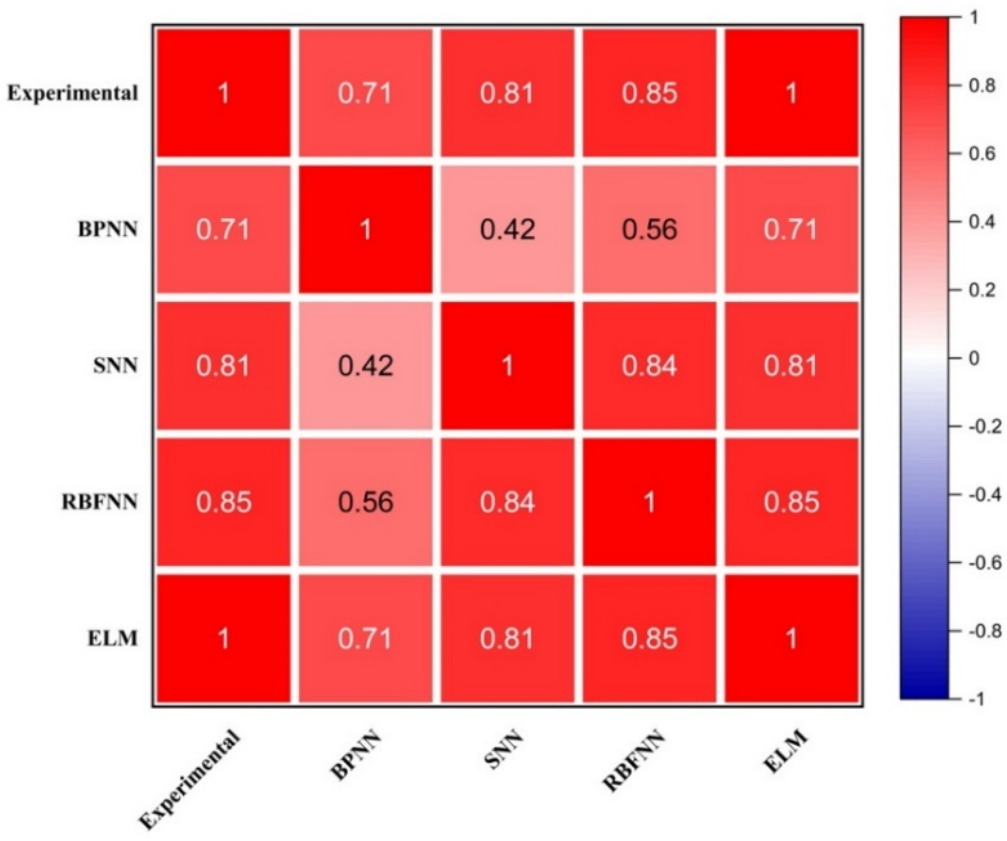

Figure 7. Correlation Matrix between Experimental data and the models for the train set

7.0 Limitations and Conclusions

The paper explores the efficacy of neural networks in modelling and predicting natural hydrogen production from geochemical reactions. Four methods were used including a method which is based on the spiking characteristics of the brain's neurons. To evaluate and compare the efficiency of the models, seven metrics were adopted (MSE, RMSE, MAE, NRMSE, R, R^2, and NSE), chosen based on the wide applicability and acceptance in assessing the performance of neural networks. Results show that the Extreme Learning machine outperforms the other models. We posit this occurrence to be attributed to the simplicity and speed of the ELM architecture. On this occasion, the handful of datasets utilised in this study did not favour the SNN.

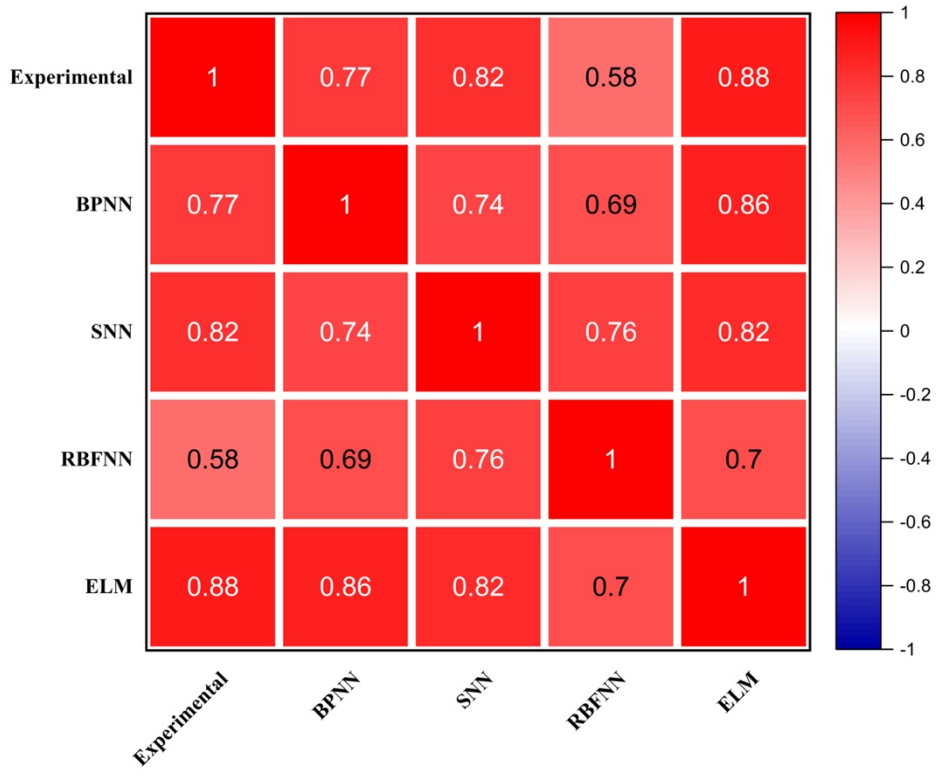

Figure 8. Correlation Matrix between Experimental data and the models for the test set

Spiking Neural Networks are more complex models that simulate the behaviour of biological neurons with spikes. While these models have shown promise in certain applications, they were not well-suited for this task, and their complexity led to challenges in training and performance optimization. The SNN in the future can be assessed on larger and more complex datasets which could potentially include a wider temperature range, mineralogical makeup of the reactants, chiefly its iron content, and even pressure to assess their effectiveness, but to this end, the work reports the effectiveness and varying abilities of neural networks, and provides researchers with information on what could work and what choice could be made when undertaking geochemical prediction studies. In all, this study has proven that neural networks can continue to make contributions to the energy transition and in modelling and predicting clean energy.

8.0 References

kjærseth, S. Andresen, G. Bang, and G. M. Heggelund, "The Paris agreement and key actors' domestic climate policy mixes: comparative patterns," *Int Environ Agreem*, vol. 21, no. 1, pp. 59–73, Mar. 2021, doi: 10.1007/s10784-021-09531-w.

[2] K. J. Kroeker *et al.*, "Impacts of ocean acidification on marine organisms: Quantifying sensitivities and interaction with warming," *Glob Chang Biol*, vol. 19, no. 6, pp. 1884–1896, Jun. 2013, doi: 10.1111/gcb.12179.

[3] "Schroders The impact of climate change on the global economy 2."

[4] J. Dorrell and K. Lee, "The cost of wind: Negative economic effects of global wind energy development," *Energies*, vol. 13, no. 14. MDPI AG, Jul. 01, 2020. doi: 10.3390/en13143667.

[5] I. Matino and V. Colla, "The role of hydrogen for a sustainable steelmaking process," *Materiaux et Techniques*, vol. 111, no. 2. EDP Sciences, 2023. doi: 10.1051/mattech/2023026.

[6] C. Likkasit, A. Maroufmashat, A. Elkamel, H. ming Ku, and M. Fowler, "Solar-aided hydrogen production methods for the integration of renewable energies into oil and gas industries," *Energy Convers Manag*, vol. 168, pp. 395–406, Jul. 2018, doi: 10.1016/j.enconman.2018.04.057.

[7] N. S. Muhammed, B. Haq, D. Al Shehri, A. Al-Ahmed, M. M. Rahman, and E. Zaman, "A review on underground hydrogen storage: Insight into geological sites, influencing factors and future outlook," *Energy Reports*, vol. 8. Elsevier Ltd, pp. 461–499, Nov. 01, 2022. doi: 10.1016/j.egyr.2021.12.002.

[8] B. Malvoisin *et al.*, "High-purity hydrogen gas from the reaction between BOF steel slag and water in the 473-673 K range," *Int J Hydrogen Energy*, vol. 38, no. 18, pp. 7382–7393, Jun. 2013, doi: 10.1016/j.ijhydene.2013.03.163.

[9] O. Maiga, E. Deville, J. Laval, A. Prinzhofer, and A. B. Diallo, "Characterization of the spontaneously recharging natural hydrogen reservoirs of Bourakebougou in Mali," *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-38977-y.

[10] S. Schwartz *et al.*, "Pressure-temperature estimates of the lizardite/antigorite transition in high pressure serpentinites," *Lithos*, vol. 178, pp. 197–210, Sep. 2013, doi: 10.1016/j.lithos.2012.11.023.

[11] F. Osselin, M. Pichavant, R. Champallier, M. Ulrich, and H. Raimbourg, "Reactive transport experiments of coupled carbonation and serpentinization in a natural serpentinite. Implication for hydrogen production and carbon geological storage," *Geochim Cosmochim Acta*, vol. 318, pp. 165–189, Feb. 2022, doi: 10.1016/j.gca.2021.11.039.

[12] M. Preiner *et al.*, "Serpentinization: Connecting geochemistry, ancient metabolism and industrial hydrogenation," *Life*, vol. 8, no. 4. MDPI AG, Dec. 01, 2018. doi: 10.3390/life8040041.

[13] B. M. Tutolo, W. E. Seyfried, and N. J. Tosca, "A seawater throttle on H 2 production in Precambrian serpentinizing systems," vol. 117, no. 26, pp. 14756–14763, 2020, doi: 10.1073/pnas.1921042117/-/DCSupplemental.

[14] H. M. Miller, L. E. Mayhew, E. T. Ellison, P. Kelemen, M. Kubo, and A. S. Templeton, "Low temperature hydrogen production during experimental hydration of partially-serpentinized dunite," *Geochim Cosmochim Acta*, vol. 209, pp. 161–183, Jul. 2017, doi: 10.1016/j.gca.2017.04.022.

[15] G. Etiope, N. Samardžić, F. Grassa, H. Hrvatović, N. Miošić, and F. Skopljak, "Methane and hydrogen in hyperalkaline groundwaters of the serpentinized Dinaride ophiolite belt, Bosnia and Herzegovina," *Applied Geochemistry*, vol. 84, pp. 286–296, Sep. 2017, doi: 10.1016/j.apgeochem.2017.07.006.

[16] R. Huang, W. Sun, X. Ding, Y. Zhao, and M. Song, "Effect of pressure on the kinetics of peridotite serpentinization," *Phys Chem Miner*, vol. 47, no. 7, Jul. 2020, doi: 10.1007/s00269-020-01101-x.

[17] R. Huang, M. Song, X. Ding, S. Zhu, W. Zhan, and W. Sun, "Influence of pyroxene and spinel on the kinetics of peridotite serpentinization," *J Geophys Res Solid Earth*, vol. 122, no. 9, pp. 7111–7126, Sep. 2017, doi: 10.1002/2017JB014231.

[18] T. M. McCollom, F. Klein, B. Moskowitz, T. S. Berquó, W. Bach, and A. S. Templeton, "Hydrogen generation and iron partitioning during experimental serpentinization of an olivine–pyroxene mixture," *Geochim Cosmochim Acta*, vol. 282, pp. 55–75, Aug. 2020, doi: 10.1016/j.gca.2020.05.016.

[19] P. L. Morrill *et al.*, “Geochemistry and geobiology of a present-day serpentinization site in California: The Cedars,” *Geochim Cosmochim Acta*, vol. 109, pp. 222–240, May 2013, doi: 10.1016/j.gca.2013.01.043.
[20] T. M. McCollom *et al.*, “Temperature trends for reaction rates, hydrogen generation, and partitioning of iron during experimental serpentinization of olivine,” *Geochim Cosmochim Acta*, vol. 181, pp. 175–200, May 2016, doi: 10.1016/j.gca.2016.03.002.
[21] T. M. McCollom and W. Bach, “Thermodynamic constraints on hydrogen generation during serpentinization of ultramafic rocks,” *Geochim Cosmochim Acta*, vol. 73, no. 3, pp. 856–875, Feb. 2009, doi: 10.1016/j.gca.2008.10.032.
[22] T. M. McCollom, F. Klein, and M. Ramba, “Hydrogen generation from serpentinization of iron-rich olivine on Mars, icy moons, and other planetary bodies,” *Icarus*, vol. 372, Jan. 2022, doi: 10.1016/j.icarus.2021.114754.
[23] N. Kumar Gupta, “International Journal of Emerging trends in Engineering and Development., Issue 2, Vol.4 (May 2012), ISSN 2249-6149.” [Online]. Available: https://www.researchgate.net/publication/266141357
[24] H. Vo Thanh, Y. Sugai, and K. Sasaki, “Application of artificial neural network for predicting the performance of CO2 enhanced oil recovery and storage in residual oil zones,” *Sci Rep*, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-73931-2.
[25] A. Gbadamosi *et al.*, “New-generation machine learning models as prediction tools for modeling interfacial tension of hydrogen-brine system,” *Int J Hydrogen Energy*, vol. 50, pp. 1326–1337, Jan. 2024, doi: 10.1016/j.ijhydene.2023.09.170.
[26] T. Zhang, Y. Zhang, K. Katterbauer, A. Al Shehri, S. Sun, and I. Hoteit, “Deep learning–assisted phase equilibrium analysis for producing natural hydrogen,” *Int J Hydrogen Energy*, vol. 50, pp. 473–486, Jan. 2024, doi: 10.1016/j.ijhydene.2023.09.097.
[27] G. A. Kale and C. Karakuzu, “Multilayer extreme learning machines and their modeling performance on dynamical systems[Formula presented],” *Appl Soft Comput*, vol. 122, Jun. 2022, doi: 10.1016/j.asoc.2022.108861.
[28] H. Yu, T. Xie, S. Paszczyñski, and B. M. Wilamowski, “Advantages of radial basis function networks for dynamic system design,” in *IEEE Transactions on Industrial Electronics*, Dec. 2011, pp. 5438–5450. doi: 10.1109/TIE.2011.2164773.
[29] C. Mahata, S. Ray, and D. Das, “Optimization of dark fermentative hydrogen production from organic wastes using acidogenic mixed consortia,” *Energy Convers Manag*, vol. 219, Sep. 2020, doi: 10.1016/j.enconman.2020.113047.
[30] S. Ascher, W. Sloan, I. Watson, and S. You, “A comprehensive artificial neural network model for gasification process prediction,” *Appl Energy*, vol. 320, Aug. 2022, doi: 10.1016/j.apenergy.2022.119289.
[31] H. M. U. Ayub, M. Rafiq, M. A. Qyyum, G. Rafiq, G. S. Choi, and M. Lee, “Prediction of Process Parameters for the Integrated Biomass Gasification Power Plant Using Artificial Neural Network,” *Front Energy Res*, vol. 10, Jun. 2022, doi: 10.3389/fenrg.2022.894875.
[32] M. K. Yogeswari, K. Dharmalingam, and P. Mullai, “Implementation of artificial neural network model for continuous hydrogen production using confectionery wastewater,” *J Environ Manage*, vol. 252, Dec. 2019, doi: 10.1016/j.jenvman.2019.109684.
[33] M. Nasr, A. Tawfik, S. Ookawara, and M. Suzuki, “PREDICTION OF HYDROGEN PRODUCTION USING ARTIFICIAL NEURAL NETWORK,” 2013.
[34] P. Jha, E. B. G. Kana, and S. Schmidt, “Can artificial neural network and response surface methodology reliably predict hydrogen production and COD removal in an UASB bioreactor?,” *Int J Hydrogen Energy*, vol. 42, no. 30, pp. 18875–18883, Jul. 2017, doi: 10.1016/j.ijhydene.2017.06.063.
[35] P. Karthic, S. Joseph, N. Arun, and S. Kumaravel, “Optimization of biohydrogen production by Enterobacter species using artificial neural network and response surface methodology,” *Journal of Renewable and Sustainable Energy*, vol. 5, no. 3, May 2013, doi: 10.1063/1.4803746.
[36] F. Azaman *et al.*, “Application of artificial neural network and response surface methodology for modelling of hydrogen production using nickel loaded zeolite,” *J Teknol*, vol. 77, no. 1, pp. 109–118, Nov. 2015, doi: 10.11113/jt.v77.4265.
[37] E. L. M. Cárdenas, A. D. Zapata-Zapata, and D. Kim, “Modeling dark fermentation of coffee mucilage wastes for hydrogen production: Artificial neural network model vs. fuzzy logic model,” *Energies (Basel)*, vol. 13, no. 7, 2020, doi: 10.3390/en13071663.
[38] Q. Zhao, H. Wang, and C. Chen, “Underground hydrogen storage: A recovery prediction using pore network modeling and machine learning,” *Fuel*, vol. 357, Feb. 2024, doi: 10.1016/j.fuel.2023.130051.
[39] S. Atuahene, Y. Bao, Y. Yevenyo Ziggah, and P. Semwaah Gyan, “Artificial Neural Network based Artificial Intelligent Algorithms for Accurate Monthly Load Forecasting of Power Consumption,” 2019.
[40] K. Baddari, T. Aïfa, N. Djarfour, and J. Ferahtia, “Application of a radial basis function artificial neural network to seismic data inversion,” *Comput Geosci*, vol. 35, no. 12, pp. 2338–2344, 2009, doi: 10.1016/j.cageo.2009.03.006.
[41] B. Konakoglu and L. Cakir, “Generalized Regression Neural Network for Coordinate Transformation,” 2018. [Online]. Available: https://www.researchgate.net/publication/329365471
[42] A. Jaiswal, S. Roy, G. Srinivasan, and K. Roy, “Proposal for a Leaky-Integrate-Fire Spiking Neuron based on Magneto-Electric Switching of Ferro-magnets,” Sep. 2016, doi: 10.1109/TED.2017.2671353.
[43] S. Agatonovic-Kustrin and R. Beresford, “Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research,” 2000. [Online]. Available: www.elsevier.com/locate/jpba
[44] L. Yitian and R. R. Gu, “Modeling flow and sediment transport in a river system using an artificial neural network,” *Environ Manage*, vol. 31, no. 1, pp. 122–134, Jan. 2003, doi: 10.1007/s00267-002-2862-9.
[45] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” Dec. 2014, [Online]. Available: http://arxiv.org/abs/1412.6980
[46] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks.”
[47] J. Wang, S. Lu, S. H. Wang, and Y. D. Zhang, “A review on extreme learning machine,” *Multimed Tools Appl*, vol. 81, no. 29, pp. 41611–41660, Dec. 2022, doi: 10.1007/s11042-021-11007-7.
[48] G. Goos *et al.*, “LNCS 3973 - Advances in Neural Networks - ISNN 2006,” 2006.
[49] X. C. Yang, X. R. Yan, and C. F. Song, “Pressure Prediction of Coal Slurry Transportation Pipeline Based on Particle Swarm Optimization Kernel Function Extreme Learning Machine,” *Math Probl Eng*, vol. 2015, 2015, doi: 10.1155/2015/542182.